

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
14 March 2002 (14.03.2002)

PCT

(10) International Publication Number  
WO 02/20786 A2

(51) International Patent Classification<sup>7</sup>: C12N 15/12,  
C07K 14/705, C12N 9/12, 15/63, 5/10, C07K 16/28,  
A61K 38/17, A01K 67/027, A61P 35/00

(72) Inventors; and

(75) Inventors/Applicants (for US only): SCHIAVO, Giampietro [IT/GB]; Molecular Neuropathobiology, Imperial Cancer Research Fund, 44 Lincoln's Inn Fields, London WC2A 3PX (GB). IGLESIAS, Teresa [ES/ES]; Instituto de Investigaciones Biomédicas, CSIC, Calle Arturo Duperier, 4, E-28029 Madrid (ES).

(21) International Application Number: PCT/GB01/03977

(22) International Filing Date:  
6 September 2001 (06.09.2001)

(25) Filing Language: English

(74) Agent: MILES, John, S.; Eric Potter Clarkson, Park View House, 58 The Ropewalk, Nottingham NG1 5DD (GB).

(26) Publication Language: English

(30) Priority Data:  
60/230,449 6 September 2000 (06.09.2000) US

(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

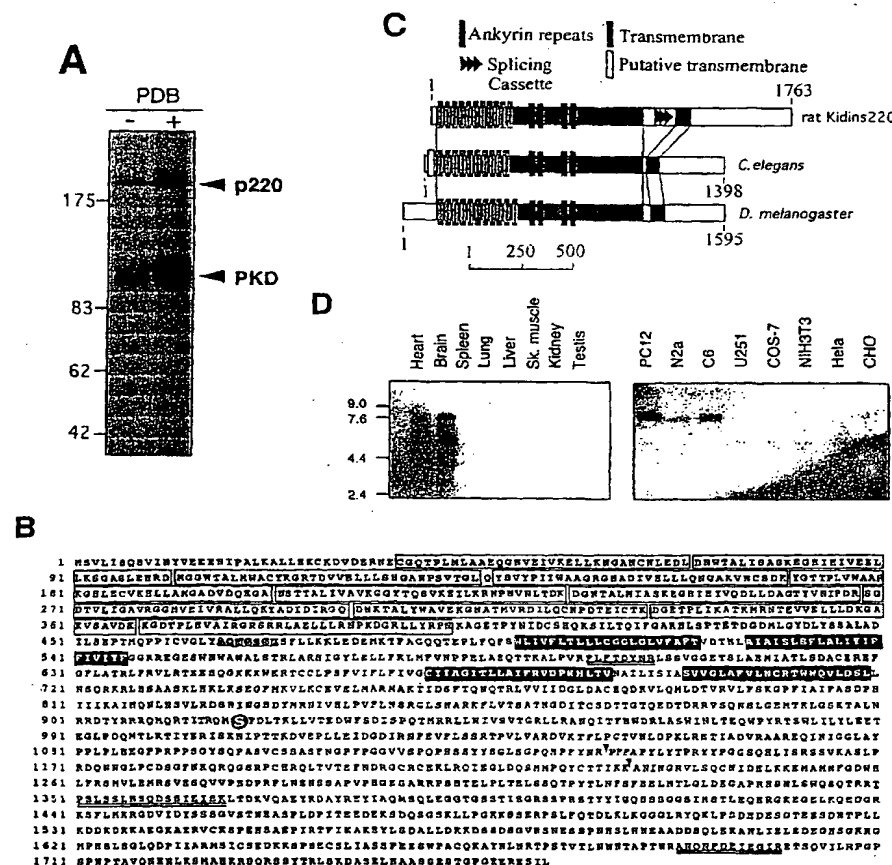
(71) Applicant (for all designated States except US): IMPE-  
RIAL CANCER RESEARCH TECHNOLOGY LIM-  
ITED [GB/GB]; Sardinia House, Sardinia Street, London  
WC2A 3NL (GB).

[Continued on next page]

(54) Title: NOVEL COMPOUNDS AND METHODS



WO 02/20786 A2



(57) Abstract: A polynucleotide comprising a nucleotide sequence which encodes the polypeptide whose amino acid sequence is given in Figure 1(b) or Figure 11) or a fragment or variant or fusion thereof or a fusion of said fragment or variant, or its complement, the polypeptide encoded by said polynucleotide and fragments or variants or fusions thereof, antibodies specifically binding the said polypeptide or fragments or variants or fusions thereof. The invention further comprises methods of producing said polypeptide and antibodies and uses thereof.



(84) **Designated States (regional):** ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

**Published:**

— without international search report and to be republished upon receipt of that report

*For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

## NOVEL COMPOUNDS AND METHODS

The present invention relates to novel compounds and methods and, in particular, it relates to a physiological substrate for PKD/PKC $\mu$  and  
5 methods of use.

PKD/PKC $\mu$  is an abundant and widespread protein kinase with unique domain composition and enzymatic properties, until now awaiting the identification of any physiological function(s) and cellular substrate(s).

10

Plasma membrane components play a crucial role in transducing extracellular signals into the cell. Stimulation of several plasma membrane receptors, by hormones, growth factors and cytokines, causes the rapid hydrolysis of phosphoinositides which results in an increase of the lipid-  
15 derived second messenger diacylglycerol (DAG; Exton, 1990). DAG causes the activation of protein kinase C (PKC; Nishizuka, 1992; Divecha and Irvine, 1995), which in turn affects several cellular processes including cell growth and differentiation, tumor promotion, changes in cell morphology, neuronal development and endocrine and exocrine secretion (Kikkawa and  
20 Nishizuka, 1986; Rozengurt, 1986; Herschmann, 1991; Nishizuka, 1992). PKC isoforms are divided into three major groups on the basis of their primary structure and biochemical properties (Newton, 1997; Mellor and Parker, 1998; Ron and Kazanietz, 1999). The conventional PKCs ( $\alpha$ ,  $\beta$ 1,  $\beta$ 2 and  $\gamma$ ) contain a regulatory cysteine-rich domain that binds DAG and  
25 phorbol esters, and a calcium binding domain, responsible for the calcium-dependent modulation of their kinase activity. The novel PKCs ( $\delta$ ,  $\epsilon$ ,  $\eta$  and  $\tau$ ) are sensitive to DAG or phorbol esters, but are calcium independent. The third group is known as the atypical PKCs ( $\zeta$  and  $\lambda$ 1), which are unresponsive to calcium, DAG or phorbol esters.

Protein Kinase D (PKD, also known as PKC $\mu$ ) is a serine/threonine kinase  
distantly related to the PKC family, that contains a conserved DAG/phorbol  
ester-binding cysteine-rich domain like the classical and novel PKC  
5 isoforms, but is characterised by unique enzymological and structural  
properties (Johannes *et al.*, 1994; Valverde *et al.*, 1994; Rozengurt *et al.*,  
1995; Iglesias *et al.*, 1998a). Unlike all other PKCs, the amino-terminal  
region of PKD lacks the typical auto-inhibitory pseudo-substrate site and  
contains a pleckstrin homology (PH) domain, unique within the PKC  
10 family, that regulates its activity (Iglesias and Rozengurt, 1998b). In intact  
cells, DAG, phorbol esters, mitogenic neuropeptides and growth factors  
induce activation of PKD through a PKC-dependent signalling pathway  
(Van Lint *et al.*, 1995; Zugaza *et al.*, 1996; Zugaza *et al.*, 1997), involving  
the phosphorylation of two activation loop sites in the carboxy-terminal  
15 catalytic domain of PKD (Iglesias *et al.*, 1998c; Waldron *et al.*, 1999). The  
increased activity of phosphorylated PKD is maintained during cell  
disruption and immunoprecipitation. This characteristic further  
distinguishes PKD from other PKC members, whose *in vitro* activity is  
strictly dependent on the presence of lipid cofactors (Zugaza *et al.*, 1996).  
20 All these unique features and the divergence of the catalytic domain of PKD  
from the highly conserved kinase sub-domains of the other PKC  
isoenzymes, suggests the involvement of PKD in distinct cellular  
function(s).

25 Selective protein-protein interaction and/or the phosphorylation of unique  
downstream effectors are likely to be the basis of isotype-specific functions  
*in vivo*. Each PKC family member shows an optimal substrate consensus  
sequence and PKD, in particular, has a substrate specificity very different  
from other PKC isoforms (Nishikawa *et al.*, 1997). Consistent with this, it



has been reported previously that PKD failed to phosphorylate several substrates which are actively phosphorylated by other members of the PKC family (Valverde *et al.*, 1994; Van Lint *et al.*, 1995). The substrates for PKD known so far are two peptides derived from glycogen synthase, the synthetic peptides syntide-2 and GS (Valverde *et al.*, 1994; Dieterich *et al.*, 1996). Several proteins, such as phosphatidylinositol-kinases (Nishikawa *et al.*, 1998), 14.3.3 proteins (Hausser *et al.*, 1999), B-cell receptor (BCR) complex, the tyrosine kinase Syk, the phospholipase C $\gamma$ 1 (PLC $\gamma$ 1; Sidorenko *et al.*, 1996) and Bruton's tyrosine kinase (Btk; Johannes *et al.*, 1999) have been reported to interact with PKD. Recently, we demonstrated the direct interaction of PKD with two members of the novel PKCs, PKC $\epsilon$  and PKC $\eta$  (Waldron *et al.*, 1999). However, none of these interacting proteins have been shown to be PKD substrates *in vivo*. These unique structural and functional features of PKD suggest that this enzyme may be involved in a still poorly understood signalling pathway with distinct cellular downstream effectors.

To gather insights into PKD cellular function(s), we have used specific anti-PKD/PKC $\mu$  antibodies to identify putative partners of PKD.

Following this strategy has surprisingly led us to purify, clone and characterise a novel PKD-interacting protein, which we have called Kidins220, that co-immunoprecipitates with PKD.

Here we describe the characterisation of the novel protein, termed Kidins220 (for Kinase D interacting substrate of 220 kDa), which binds to PKD and acts as a specific substrate. By using confocal microscopy as well as biochemical fractionation, we demonstrate that Kidins220 is a novel integral plasma membrane protein selectively expressed in brain and

neurocrine cells. Kidins220 segregates in lipid rafts and, during neurite formation, concentrates at the growth cones, both on the plasma membrane and on intracellular vesicular compartments. Kidins220 not only interacts with PKD, but is also the first protein that has been shown to be phosphorylated by PKD. Moreover, a peptide corresponding to the consensus PKD-phosphorylation motif, present within the Kidins220 sequence, behaves as one of the best peptide substrates identified so far. Finally, by transient expression of PKD mutants, we show that PKD activity is essential for the stability of Kidins220 *in vivo*. Inhibition of PKD leads to the aggregation and degradation of Kidins220, whereas an increase of the PKD kinase activity promotes Kidins220 accumulation in the cell. These findings indicate that PKD is an essential modulator of Kidins220 trafficking and turnover and suggest a striking new role for the PKD pathway in the regulation of the stability of its substrate(s).

15

Two sequences corresponding to different fragments of the human Kidins220 homologue have been deposited with GenBank (AB033076 and AL133620; Nagase *et al.*, 1999). However, these fragments are incomplete sequences. Also, there is no disclosure in these GenBank entries or in Nagase *et al.*, 1999 of the rat cDNA, the complete Kidins220 protein, or the function of the Kidins220 protein.

20

Subsequent to the filing of our priority application, Kong *et al* (2001) *J. Neurosci.* 21:176-185 (incorporated herein by reference) have cloned a polynucleotide which encodes a polypeptide with sequence identity to the human and rat Kidins220 polypeptide. The polypeptide, designated ARMS (for ankyrin, repeat rich, and membrane spanning), is reported to be a downstream target of the neurotrophin and ephrin receptor tyrosine kinases Trk and Eph respectively, and is shown to physically associate with the

25

TrkA and p75 neurotrophin receptors. Although the Kidins220/ARMS polypeptide is also shown to be phosphorylated on a tyrosine residue in Kong *et al* (2001), there is no indication of which tyrosine residue, and no suggestion of which kinase is acting on the protein. On the basis of its  
5 integral membrane structure, Kong *et al* (2001) propose that the Kidins220/ARMS polypeptide may function as an ion channel. It is further suggested that the multiple protein interaction domains may indicate a function in recruiting proteins to Trk receptor tyrosine kinases.

10 We have also shown in this work that seven additional forms of the rat Kidins220 protein exist. These forms appear to arise due to differential splicing of the transcript.

Knowledge of the novel protein Kidins220, its cDNA and amino acid  
15 sequences, and as a physiological target of PKD, now allows the study of the role of PKD in the regulation of protein stability and provides a target mechanism for manipulation in the treatment of pathogenic processes, such as neurodegenerative diseases or cancer.

20 A first aspect of the invention provides a polynucleotide comprising a nucleotide sequence which encodes the polypeptide whose amino acid sequence is given in Figure 1b or Figure 11, or a fragment or variant or fusion thereof or a fusion of said fragment or variant or a complement thereof, provided that the polynucleotide is not any one of the clones  
25 corresponding to the GenBank Accession Nos AL133620, AB033076, AC007175, F36H1 or any Accession No. listed in Figures 14a or 14b.

The polynucleotide may be DNA or RNA but it is preferred if it is DNA. The polynucleotide may or may not contain introns. It is preferred if it does

not contain introns, and it is particularly preferred if the polynucleotide is a cDNA.

Figure 1b shows the amino acid sequence of the rat Kidins220 polypeptide.

- 5 It is preferred if the polynucleotide encodes a polypeptide comprising the amino acid sequence given in Figure 1b, or variants or fragments or fusions thereof, or fusions of said variants or fragments. It is particularly preferred if the polynucleotide encodes a polypeptide consisting of the amino acid sequence given in Figure 1b.

10

Preferably the polynucleotide is not the partial cDNA clone identified in WO 01/42285 as SEQ ID No 30 or a fragment thereof consisting of nucleotides 1-285, or 813-930 or 1145-3381 as numbered in SEQ ID No 30.

- 15 Throughout the specification where the term Kidins220 is used, it is intended to include a polypeptide whose amino acid sequence is given in Figure 1b or the human homologue of this sequence, unless the context suggests otherwise. The amino acid sequence of the human homologue (human Kidins220) is given in Figure 11.

20

- A polynucleotide of the invention is not a polynucleotide which is any one of the DNA clones corresponding to the GenBank Accession Nos AL133620, AB033076, AC007175, F36H1 or any Accession No. listed in Figures 14a or 14b. Preferably the polynucleotide of the invention is not the  
25 partial cDNA clone identified in WO 01/42285 as SEQ ID No 30 or a fragment thereof consisting of nucleotides 1-285, or 813-930 or 1145-3381 as numbered in SEQ ID No 30. GenBank Accession No AL133620 is a partial cDNA cloned within the vector pSport1. The cDNA sequence appears to encode a C-terminal portion of the human Kidins220 homologue.

GenBank Accession No AB033076 is a partial cDNA cloned into the pBluescriptII SK plus vector. The cDNA sequence appears to encode an N-terminal portion of the human Kidins220 homologue. GenBank Accession No AC007175 is a *Drosophila* chromosome 2R DNA sequence contained within a BAC clone. Accession No F36H1 is a cosmid containing *C. elegans* genome sequence. Accession No AA998517 is an Expressed Sequence Tag (EST) cloned into the pT7T3D-Pac vector. Accession No N83544 is an EST cloned into a Lambda ZAP Express vector and Accession No AA685317 is an EST cloned into a pBluescript SK- vector.

10

The partial cDNA clone identified as SEQ ID No 30 in WO 01/42285 is a human DNA sequence contained within the plasmid pINCY (Incyte Genomics, Palo Alto, CA).

15 For the avoidance of doubt, however, the polynucleotides of the invention include polynucleotides which encode the amino sequences described in the GenBank Accession Nos listed in Figure 14b since these sequences are believed to be derived from the human Kidins220 protein but as noted above do not include the specific DNA clones corresponding to the given

20 GenBank Accession Nos. It is particularly preferred if the polynucleotides which encode a polypeptide which contains the amino acid sequences described in the GenBank Accession Nos listed in Figure 14b encode the human Kidins220 polypeptide. It is particularly preferred if the polynucleotides of the invention contain the nucleotide sequence given in

25 GenBank Accession Nos listed in Figure 14b. A sequence of human Kidins220 cDNA is shown in Figure 13. The sequence shows a greater than 91% homology with the sequence encoding rat Kidins220 polypeptide shown in Figure 1b and as such is within the scope of this invention.

In a preferred embodiment, the polynucleotide of the first aspect of the invention comprises a nucleic acid sequence given in Figure 7 or 17. It is particularly preferred if the polynucleotide comprises or consists of the sequence corresponding to the open reading frame (ORF) in Figure 7 or 17  
5 as this encodes the rat Kidins220 polypeptide. It is also particularly preferred if the polynucleotide comprises or consists of the sequence corresponding to the ORF in Figure 13 as this encodes the human Kidins220 polypeptide.

10 Polynucleotides of the invention also include polynucleotides which encode any polypeptides encoded by the specific polynucleotides described such as the polynucleotides described in Figures 7, 17 and 13.

In another embodiment of the first aspect of the invention the  
15 polynucleotide encodes a polypeptide whose amino acid sequence is as described in Figure 8. These polypeptides appear to be produced by splice variants of the Kidins220 mRNA/cDNA.

In particular, amino acid residues between the arrows in Figure 1b (residue  
20 numbers 1140 and 1232) are believed to constitute a "splicing cassette" wherein the "cassette" shown in Figure 1b is replaced in the variants with an alternative cassette whose amino acid sequence is chosen among those reported between the square brackets in Figure 8.

25 It is understood that the polynucleotide which encodes the polypeptide whose amino acid sequence is given in Figure 1b or Figure 11 (or a variant or fragment thereof) may vary in sequence due to degeneracy of the genetic code. For example, different codons can be substituted which code for the same amino acid(s) as the original codons. Alternatively, the substitute

codons may code for a different amino acid that will not affect the activity or immunogenicity of the polypeptide or which may improve or otherwise modulate its activity or immunogenicity. For example, site-directed mutagenesis or other techniques can be employed to create single or multiple mutations, such as replacements, insertions, deletions, and transpositions, as described in Botstein and Shortle, "Strategies and Applications of *In Vitro* Mutagenesis," *Science*, **229**: 193-210 (1985), which is incorporated herein by reference. Since such modified polynucleotides can be obtained by the application of known techniques to the teachings contained herein, such modified polynucleotides are within the scope of the claimed invention.

Moreover, it will be recognised by those skilled in the art that the polynucleotide sequence, (or fragments thereof) of the invention can be used to obtain other polynucleotide sequences (which may be considered to be variants) that hybridise with it under conditions of high stringency. Such polynucleotides includes any genomic DNA. The polynucleotide of the invention includes polynucleotides which hybridise to the DNA whose sequence is given in Figure 7 or Figure 13 under stringent conditions, preferably at high stringency. Accordingly, the polynucleotide of the invention includes polynucleotides that show at least 55%, preferably at least 60%, and more preferably at least 65%, 70% or 75% or 80% or 85%, and most preferably at least 90% or 95% or 98% identity with the polynucleotide whose nucleotide sequence is shown in Figure 7 or Figure 13, provided that such homologous polynucleotide encodes a polypeptide whose amino acid sequence is given in Figure 1b or Figure 11 or a variant or fragment or fusion thereof and is not any one of GenBank Accession Nos AL133620, AB033076, AC007175, F36H1 or any GenBank Accession No. listed in Figures 14a or 14b. A preferred polynucleotide variant is one

which has the nucleotide sequence shown in Figure 17, or one which encodes a polypeptide with the amino acid sequence shown in Figure 18.

Per cent identity between nucleotide sequences can be determined by, for example, the Blast 2 sequence program (T.A. Tatusova and T.L. Madden  
5 FEMS Microbiol. Lett. (1999) 174:247-250) at the NCBI site (<http://www.ncbi.nlm.nih.gov/gorf/bl2.html>) using as parameters the scoring matrix BLOSUM62, opening gap penalty -5, extending gap penalty -2.

10

It is believed that Kidins220 is found in mammals other than rat and humans. The present invention therefore includes polynucleotides which encode Kidins220 from other mammalian species including human, mouse, cow, pig, sheep, rabbit and so on, but which polynucleotide is not any one  
15 of GenBank Accession Nos AL133620, AB033076, AC007175, F36H1 or any GenBank Accession No. listed in Figures 14a or 14b or the partial cDNA clone identified in WO 01/42285 as SEQ ID No 30 or fragments thereof consisting of nucleotides 1-285, 813-930 or 1145-3381 as numbered in SEQ ID No 30.

20

The polynucleotide of the invention has at least 10 nucleotides, and preferably at least 15 nucleotides and more preferably at least 30 nucleotides. In a further preference, the polynucleotide is more than 50 nucleotides, more preferably at least 100 nucleotides, and still more  
25 preferably the polynucleotide is at least 500 nucleotides. The polynucleotide may be more than 1kb, and may comprise more than 5kb.

In one embodiment, the polynucleotide encodes a polypeptide whose amino acid sequence comprises the sequence CAASSESTGFGEERSIL or



RQMQRRTITRQMSFDLTK or RQMQRRTITRQMAFDLTK or a polypeptide which comprises the amino acid sequence (I/L)XRQM(S/A)J, where X is any amino acid, J represents a hydrophobic residue such as isoleucine, leucine, phenylalanine, tryptophan, tyrosine or valine, (I/L) represents isoleucine or leucine, and (S/A) represents serine or alanine. Preferably, the polynucleotide encodes a polypeptide which comprises the amino acid sequence (I/L)XRQM(S/A)J and further amino acids to the N-terminus and C-terminus of this motif present in rat or human Kidins220. Typically the polynucleotide encodes a polypeptide of at least 10 amino acids, preferably at least 20 amino acids and more preferably at least 50 or 100 amino acids. Such polynucleotides are useful for expressing large amounts of these polypeptides which are particularly useful in procedures such as raising antibodies or competition assays. As described below, the peptide CAASSESTGFGGEERSIL is useful for raising antibodies which are selective for Kidins220, and the sequence RQMQRRTITRQMSFDLTK contains a phosphorylation site for Protein Kinase D (underlined).

It is further preferred if the polynucleotide comprises the open reading frame given in Figure 7 or Figure 13 or Figure 17. Such a polynucleotide is useful for molecular manipulations where the full length Kidins220 cDNA is to be used, for example, in generating a DNA construct suitable for expressing the full length Kidins220 polypeptide.

The invention also includes a polynucleotide which is able to selectively hybridise to a polynucleotide which encodes the polypeptide whose amino acid sequence is given in Figure 1b or Figure 11, or a fragment or variant or fusion thereof, or a fusion of said variant or fragment. A preferred variant to which the polynucleotide is able to selectively hybridise is that shown in Figure 17. Preferably, said polynucleotide is at least 10 nucleotides, more

preferably at least 15 nucleotides and still more preferably at least 30 nucleotides in length. The said polynucleotide may be longer than 100 nucleotides and may be longer than 200 nucleotides, but preferably the said polynucleotide is not longer than 250 nucleotides. Such polynucleotides are  
5 useful in procedures as a detection tool to demonstrate the presence of the polynucleotide in a sample. Such a sample may be a sample of DNA, such as a bacterial colony, fixed on a membrane or filter, or may be a sample from a patient, particularly a patient in whom cancer has been found or is suspected.

10

By "selectively hybridise" we mean that the polynucleotide hybridises under conditions of high stringency. DNA-DNA, DNA-RNA and RNA-RNA hybridisation may be performed in aqueous solution containing between 0.1X SSC and 6X SSC and at temperatures of between 55°C and  
15 70°C. It is well known in the art that the higher the temperature or the lower the SSC concentration the more stringent the hybridisation conditions. By "high stringency" we mean 2X SSC and 65°C. 1X SSC is 0.15M NaCl/0.015M sodium citrate. Polynucleotides which hybridise at high stringency are included within the scope of the claimed invention.

20

In another embodiment, the polynucleotide can be used as a primer in the polymerase chain reaction (PCR), and in this capacity a polynucleotide of between 15 and 30 nucleotides is preferred. A polynucleotide of between 20 and 100 nucleotides is preferred when the fragment is to be used as a  
25 mutagenic PCR primer. It is particularly preferred if the PCR primer (when not being used to mutate a nucleic acid) contains about 15 to 30 contiguous nucleotides (ie perfect matches) from the nucleotide sequence given in Figure 7 or 17 or the human cDNA sequence shown in Figure 13. Clearly,

if the PCR primers are used for mutagenesis, differences compared to the sequence will be present.

Primers which are suitable for use in a polymerase chain reaction (PCR; 5 Saiki *et al* (1988) *Science* **239**, 487-491) are preferred. Suitable PCR primers may have the following properties:

It is well known that the sequence at the 5' end of the oligonucleotide need not match the target sequence to be amplified.

10

It is usual that the PCR primers do not contain any complementary structures with each other longer than 2 bases, especially at their 3' ends, as this feature may promote the formation of an artifactual product called "primer dimer". When the 3' ends of the two primers hybridize, they form 15 a "primed template" complex, and primer extension results in a short duplex product called "primer dimer".

Internal secondary structure should be avoided in primers. For symmetric PCR, a 40-60% G+C content is often recommended for both primers, with 20 no long stretches of any one base. The classical melting temperature calculations used in conjunction with DNA probe hybridization studies often predict that a given primer should anneal at a specific temperature or that the 72°C extension temperature will dissociate the primer/template hybrid prematurely. In practice, the hybrids are more effective in the PCR 25 process than generally predicted by simple  $T_m$  calculations.

Optimum annealing temperatures may be determined empirically and may be higher than predicted. *Taq* DNA polymerase does have activity in the 37-55°C region, so primer extension will occur during the annealing step

and the hybrid will be stabilized. The concentrations of the primers are equal in conventional (symmetric) PCR and, typically, within 0.1- to 1nM range.

- 5 When a pair of suitable nucleic acids of the invention are used in a PCR it is convenient to detect the product by gel electrophoresis and ethidium bromide staining. As an alternative to detecting the product of DNA amplification using agarose gel electrophoresis and ethidium bromide staining of the DNA, it is convenient to use a labelled oligonucleotide
- 10 capable of hybridising to the amplified DNA as a probe. When the amplification is by a PCR the oligonucleotide probe hybridises to the interprimer sequence as defined by the two primers. The probe may be labelled with a radionuclide such as  $^{32}\text{P}$ ,  $^{33}\text{P}$  and  $^{35}\text{S}$  using standard techniques, or may be labelled with a fluorescent dye. When the
- 15 oligonucleotide probe is fluorescently labelled, the amplified DNA product may be detected in solution (see for example Balaguer *et al* (1991) "Quantification of DNA sequences obtained by polymerase chain reaction using a bioluminescence adsorbent" *Anal. Biochem.* **195**, 105-110 and Dilesare *et al* (1993) "A high-sensitivity electrochemiluminescence-based
- 20 detection system for automated PCR product quantitation" *BioTechniques* **15**, 152-157.

- PCR products can also be detected using a probe which may have a fluorophore-quencher pair or may be attached to a solid support or may
- 25 have a biotin tag or they may be detected using a combination of a capture probe and a detector probe.

Fluorophore-quencher pairs are particularly suited to quantitative measurements of PCR reactions (eg RT-PCR). Fluorescence polarisation using a suitable probe may also be used to detect PCR products.

- 5 Oligonucleotide primers can be synthesised using methods well known in the art, for example using solid-phase phosphoramidite chemistry.

As has been referred to above, and is discussed in more detail in the Examples, splice variants of Kidins220 exist. A preferred polynucleotide of the invention includes those polynucleotides which can distinguish the usual Kidins220 cDNA/mRNA (eg that encodes the Kidins220 amino acid sequence shown in Figure 1b) from the splice variants (eg those whose amino acid differences are shown in Figure 8) or the different splice variants from each other. Typically, the polynucleotide in this embodiment is one which hybridises to a portion of nucleic acid which encodes a portion of the usual Kidins220 amino acid sequence but does not encode a portion of the variant (ie is specific for the usual Kidins220 cDNA/mRNA form). Alternatively, the polynucleotide may be one which hybridises to a portion of nucleic acid which encodes a portion of a specific splice variant Kidins220 amino acid sequence but does not encode a portion of the usual Kidins220 or of another splice variant (ie is specific for a particular splice variant). Suitable polynucleotides may readily be made by the skilled person by reference to Figures 1b, 7, 8 and 17.

- 25 It will also be appreciated that PCR primers may be made by reference to those Figures which can be used to distinguish the splice variants. For example, the PCR primers may be ones which hybridise to nucleotide sequences common to the usual Kidins220 mRNA/cDNA and the splice variants, and the presence or absence of the usual Kidins220 or particular

splice variants determined by reference to the size of DNA amplified. For example, PCR primer which hybridise to nucleic acid encoding portions of Kidins220 either side of the splice cassette referred to above may be used.

- 5 Methods currently used for detection of Single Nucleotide Polymorphisms can also be used to distinguish splice variants, such as Invader Assay™ (Third Wave Technologies, Inc.), assays based on detection of primer extension such as SNP-IT™ (Orchid Biosciences, Inc.) and Rolling Circle Amplification (Lizardi *et al* (1998) *Nat. Genet.* 19(3), 225-32).

10

- A polynucleotide of the first aspect of the invention may encode a variant of the polypeptide whose amino acid sequence is given in Figure 1b or Figure 11. A "variant" of the polypeptide includes natural variants, including allelic variants and naturally-occurring mutant forms, "splice variants" (including those whose regions of variation in amino acid sequence are shown in Figure 8), and variants with insertions, deletions and substitutions, either conservative or non-conservative, where such changes do not substantially alter the activity of the said polypeptide. In addition, the insertions and/or deletions may lead to frameshift mutations which may encode truncated (or elongated) polypeptide products, and insertions, deletions or other mutations may lead to the introduction of stop codons which encode truncate polypeptide products. A preferred variant of the polypeptide is one which comprises the sequence shown in Figure 18.

- 25 By "conservative substitution" is intended combinations such as Gly, Ala; Val, Ile, Leu; Asp, Glu; Asn, Gln; Ser, Thr; Lys, Arg; and Phe, Tyr.

By "non-conservative substitution" we include other substitutions, such as those where the substituted residue mimics a particular modification of the

replaced residue, for example a phosphorylated serine may be replaced by aspartate or glutamate due to the similarity of the aspartate or glutamate side chain to a phosphorylated serine (ie they carry a negative charge at neutral pH).

5

Further non-conservative substitutions which are included in the term "variants" are point mutations which alter one, sometimes two, and usually no more than three amino acids. Such mutations are well known in the art of biochemistry and are usually designed to insert or remove a defined characteristic of the polypeptide. Examples of useful mutations in this invention are mutation of serine 919 (or its equivalent, such as serine 918 of Figure 18) to a non-phosphorylatable residue, such as alanine, to provide a form of the polypeptide which is not a target for Protein Kinase D but may still compete for the active site of the kinase. Another type of non-conservative mutation is the alteration or addition of a residue to a cysteine or lysine residue which can then be used with maleimide or succinimide cross-linking reagents to covalently conjugate the polypeptide to another moiety. Non-glycosylated proteins may be mutated to convert an asparagine to the recognition motif N-X-S/T for N-linked glycosylation. Such a modification may be useful to create a tag for purification of the polypeptide using Concanavalin A-linked beads.

Other useful mutations include the replacement of one or more lysine residues in the motifs SKLLPGKKSS (residues 1481-1490) and SICSEDKKS (residues 1637-1645) which are putative ubiquitinylation sites. Other useful mutations also include the replacement of one or more residues (especially lysine) in the motif AQWGSGKS (residues 467-474) which is a putative ATP binding motif and the replacement of one or more tyrosine residues in the motifs KAGETPY (residues 403-409) and

KRGDVIDY (residues 1446-1453) which are putative tyrosine phosphorylation sites.

Such variants may be made using the methods of protein engineering and site-directed mutagenesis well known in the art.

Variants of the Kidins220 polypeptide include polypeptides with at least 65% identity to the amino acid sequence given in Figure 1b or Figure 11, preferably at least 70% or 80% or 85% or 90% identity to said sequence, and more preferably at least 95% or 98% identity to said amino acid sequence. A preferred variant is one with the sequence shown in Figure 18.

Percent identity can be determined by, for example, the LALIGN program (Huang and Miller, Adv. Appl. Math. (1991)12:337-357) at the Expasy facility site ([http://www.ch.embnet.org/software/LALIGN\\_form.html](http://www.ch.embnet.org/software/LALIGN_form.html)) using as parameters the global alignment option, scoring matrix BLOSUM62, opening gap penalty -14, extending gap penalty -4. Comparison between rat and *C.elegans* Kidins220 is shown in Figure 10.

Preferably, the polynucleotide encodes a Kidins220 that has at least 30%, preferably at least 50% and more preferably at least 70% of the activity of a natural Kidins220, under the same assay conditions.

Kidins220 may be assessed in bioassays based on its function as a physiological substrate of Protein Kinase D and phosphorylation by this kinase. Kidins220 may also be assessed in bioassays based on its ability to bind ATP or analogues thereof. Kidins220 may also be assayed using immunological techniques such as ELISA and the like. Further functions of Kidins220 are as a substrate for a tyrosine kinase, preferably a tyrosine



kinase acting downstream of neurotrophin and ephrin receptor tyrosine kinases, and as a protein capable of associating with TrkA and p75 neurotrophin receptors. Such tyrosine phosphorylation and protein association is described in Kong *et al* (2001) *J. Neuroscience*, 21:176-185  
5 (incorporated herein by reference).

Further activities resulting from function of Kidins220 which may be assessed are cell differentiation, neurite elongation and growth cone progression (Lamoureux *et al* (1997) *J. Cell Science* 110:635-4; Isbister and  
10 O'Connor (1999) *J. Neurosci.* 19:2589-600; Fritsche *et al* (1999) *Mol. Cell Neurosci.* 14:389-418). The polymerisation rates of cytoskeletal components such as actin or microtubules or microtubule components may be affected by the presence or form of Kidins220, and as such may also be activities suitable for reporting Kidins220 function. The interaction of  
15 Kidins220 with the polymerised actin or microtubule components may present a further activity useful for assessing Kidins220 function.

Due to the accumulation of Kidins220 in the area adjacent to the growth cone in a vesicular compartment, Kidins 220 is believed to play a role in the  
20 opposition of new membrane during neurite elongation and/or neuronal remodelling. Thus, Kidins220 is believed to have a role in CNS and PNS development. Experiments suggest that Kidins220 is present in primary motoneurone from spinal cord with a distribution similar to that observed in differentiated PC12. A failure in Kidins220 might determine failure in  
25 neuronal differentiation or in its maintenance. A neurodegenerative profile is therefore a possible pathology associated with the impairment of Kidins220 function.

By "fragment of the polypeptide whose amino acid sequence is given in Figure 1b" or Figure 11 we include any fragment which retains activity of the full length polypeptide or which is useful in some other way, for example, for use in raising antibodies or in a binding assay, but which is not the polypeptide expressed in Nagase *et al* (1999) *DNA Res.* 6; 337-345.

As discussed below, useful fragments also include those that contain the phosphorylatable serine (serine 919). Typically, the fragments are at least five consecutive residues, preferably at least 10 consecutive residues and may be 15, 25, 50, 100, 150 or more consecutive residues.

By "fusion of the polypeptide whose amino acid sequence is given in Figure 1b or Figure 11" we include said Kidins220 polypeptide, fused to any other polypeptide. For example, the said Kidins220 may be fused to a polypeptide such as glutathione-S-transferase (GST) or protein A in order to facilitate purification of Kidins220, or it may be fused to some other polypeptide which imparts some desirable characteristics on the Kidins220 polypeptide, for example a fluorescent protein such as a green fluorescent protein (GFP). Fusions to any variant (for example, the sequence shown in Figure 18) or fragment are included in the scope of the invention.

Conveniently, a polynucleotide of the invention further comprises a detectable label. By "detectable label" is included any convenient radioactive label such as  $^{32}\text{P}$ ,  $^{33}\text{P}$  or  $^{35}\text{S}$  which can readily be incorporated into a nucleic acid molecule using well known methods; any convenient fluorescent or chemiluminescent label which can readily be incorporated into a nucleic acid is also included. In addition the term "detectable label" also includes a moiety which can be detected by virtue of binding to another moiety (such as biotin which can be detected by binding to streptavidin);

and a moiety, such as an enzyme, which can be detected by virtue of its ability to convert a colourless compound into a coloured compound, or *vice versa* (for example, alkaline phosphatase can convert colourless *o*-nitrophenylphosphate into coloured *o*-nitrophenol). The detectable label  
5 may also be a fluorophore-quencher pair as described in Tyagi & Kramer (1996) *Nature Biotechnology* 14, 303-308.

A second aspect of the invention provides an expression vector comprising a polynucleotide according to the first aspect of the invention.

10

A variety of methods have been developed to operably link polynucleotides, especially DNA, to vectors, for example, via complementary cohesive termini. For instance, complementary homopolymer tracts can be added to the DNA segment to be inserted into the vector DNA. The vector and DNA  
15 segment are then joined by hydrogen bonding between the complementary homopolymeric tails to form recombinant DNA molecules.

Synthetic linkers containing one or more restriction sites provide an alternative method of joining the DNA segment to vectors. The DNA  
20 segment, generated by endonuclease restriction digestion as described earlier, is treated with bacteriophage T4 DNA polymerase or *E.coli* DNA polymerase I, enzymes that remove protruding, 3'-single-stranded termini with their 3'-5'-exonucleolytic activities, and fill in recessed 3'-ends with their polymerising activities.

25

The combination of these activities therefore generates blunt-ended DNA segments. The blunt-ended segments are then incubated with a larger molar excess of linker molecules in the presence of an enzyme that is able to catalyse the ligation of blunt-ended DNA molecules, such as bacteriophage

T4 DNA ligase. Thus, the products of the reaction are DNA segments carrying polymeric linker sequences at their ends. These DNA segments are then cleaved with the appropriate restriction enzyme and ligated to an expression vector that has been cleaved with an enzyme that produces  
5 termini compatible with those of the DNA segment.

Synthetic linkers containing a variety of restriction endonuclease site are commercially available from a number of sources including International Biotechnologies Inc., New Haven, CN, USA.

10

A desirable way to modify the DNA encoding the polypeptide of the invention is to use PCR. This method may be used for introducing the DNA into a suitable vector, for example by engineering in suitable restriction sites, or it may be used to modify the DNA in other useful ways  
15 as is known in the art.

In this method the DNA to be enzymatically amplified is flanked by two specific primers which themselves become incorporated into the amplified DNA. The said specific primers may contain restriction endonuclease  
20 recognition sites which can be used for cloning into expression vectors using methods known in the art.

The DNA (or in the case of retroviral vectors, RNA) is then expressed in a suitable host to produce a polypeptide comprising the polypeptide of the  
25 invention. Thus, the DNA encoding the polypeptide constituting the polypeptide of the invention may be used in accordance with known techniques, appropriately modified in view of the teachings contained herein, to construct an expression vector, which is then used to transform an appropriate host cell for the expression and production of the polypeptide of

the invention. Such techniques include those disclosed in US Patent Nos. 4,440,859 issued 3 April 1984 to Rutter *et al*, 4,530,901 issued 23 July 1985 to Weissman, 4,582,800 issued 15 April 1986 to Crowl, 4,677,063 issued 30 June 1987 to Mark *et al*, 4,678,751 issued 7 July 1987 to Goeddel, 5 4,704,362 issued 3 November 1987 to Itakura *et al*, 4,710,463 issued 1 December 1987 to Murray, 4,757,006 issued 12 July 1988 to Toole, Jr. *et al*, 4,766,075 issued 23 August 1988 to Goeddel *et al* and 4,810,648 issued 7 March 1989 to Stalker, all of which are incorporated herein by reference.

10 The DNA (or in the case of retroviral vectors, RNA) encoding the polypeptide constituting the polypeptide of the invention may be joined to a wide variety of other DNA sequences for introduction into an appropriate host. The companion DNA will depend upon the nature of the host, the manner of the introduction of the DNA into the host, and whether episomal  
15 maintenance or integration is desired.

Generally, the DNA is inserted into an expression vector, such as a plasmid, in proper orientation and correct reading frame for expression. If necessary, the DNA may be linked to the appropriate transcriptional and translational  
20 regulatory control nucleotide sequences recognised by the desired host, although such controls are generally available in the expression vector. The vector is then introduced into the host through standard techniques. Generally, not all of the hosts will be transformed by the vector. Therefore, it will be necessary to select for transformed host cells. One selection  
25 technique involves incorporating into the expression vector a DNA sequence, with any necessary control elements, that codes for a selectable trait in the transformed cell, such as antibiotic resistance. Alternatively, the gene for such selectable trait can be on another vector, which is used to co-transform the desired host cell.

Host cells that have been transformed by the expression vector of the invention are then cultured for a sufficient time and under appropriate conditions known to those skilled in the art in view of the teachings disclosed herein to permit the expression of the polypeptide, which can then be recovered.

Many expression systems are known, including bacteria (for example, *E.coli* and *Bacillus subtilis*), yeasts (for example *Saccharomyces cerevisiae*), filamentous fungi (for example *Aspergillus*), plant cells, animal cells and insect cells.

The vectors typically include a prokaryotic replicon, such as the ColE1 *ori*, for propagation in a prokaryote, even if the vector is to be used for expression in other, non-prokaryotic, cell types. The vectors can also include an appropriate promoter such as a prokaryotic promoter capable of directing the expression (transcription and translation) of the genes in a bacterial host cell, such as *E.coli*, transformed therewith.

A promoter is an expression control element formed by a DNA sequence that permits binding of RNA polymerase and transcription to occur. Promoter sequences compatible with exemplary bacterial hosts are typically provided in plasmid vectors containing convenient restriction sites for insertion of a DNA segment of the present invention.

Typical prokaryotic vector plasmids are pUC18, pUC19, pBR322 and pBR329 available from Biorad Laboratories, (Richmond, CA, USA) and pTrc99A and pKK223-3 available from Pharmacia, Piscataway, NJ, USA.

A typical mammalian cell vector plasmid is pSVL available from Pharmacia, Piscataway, NJ, USA. This vector uses the SV40 late promoter to drive expression of cloned genes, the highest level of expression being found in T antigen-producing cells, such as COS-1 cells.

5

An example of an inducible mammalian expression vector is pMSG, also available from Pharmacia. This vector uses the glucocorticoid-inducible promoter of the mouse mammary tumour virus long terminal repeat to drive expression of the cloned gene.

10

Useful yeast plasmid vectors are pRS403-406 and pRS413-416 and are generally available from Stratagene Cloning Systems, La Jolla, CA 92037, USA. Plasmids pRS403, pRS404, pRS405 and pRS406 are Yeast Integrating plasmids (YIps) and incorporate the yeast selectable markers *HIS3*, *TRP1*, *LEU2* and *URA3*. Plasmids pRS413-416 are Yeast Centromere plasmids (Ycps).

15

Other vectors and expression systems are well known in the art for use with a variety of host cells.

20

In one embodiment of the second aspect of the invention, the expression vector is one suitable for use in gene therapy.

25

A third aspect of the invention provides a polynucleotide comprising the sequence shown in Figure 16 or a fragment or variant thereof provided that it is not the clone corresponding to GenBank Accession No AC012495.

In a preferred embodiment, the polynucleotide or fragment or variant of polynucleotide is capable of promoting transcription, and in an even more

preferred embodiment, the transcription is promoted in a neuronal cell-selective way.

By "promote transcription" we include the meaning that the polynucleotide  
5 is one which when operably linked to another polynucleotide allows for the transcription of that other polynucleotide in a suitable cell, such as a mammalian cell.

By "neuronal cell-selective" we include the meaning that the polynucleotide  
10 is able to promote transcription in neuronal cells (typically mammalian neuronal cells) but is less able or is substantially incapable of promoting transcription in non-neuronal cells. Conveniently, the polynucleotide can promote transcription at least 5 times more efficiently (using a standard transcription assay such as chloramphenicol acetyl transferase expression)  
15 in neuronal cells than non-neuronal cells. Preferably, the polynucleotide is at least 10 times or, more preferably 50 times, more efficient in neuronal cells than non-neuronal cells.

In one embodiment, the polynucleotide comprises at least 10 consecutive  
20 nucleotides located between positions 34609 and 36632 as shown in Figure 16. Preferably, the polynucleotide comprises at least 15, 20, 25, 30, 35, 40 or 50 or 100 or 200 or 500 consecutive nucleotides from between positions 34609 and 36632 as shown in Figure 16.

25 It is preferred if the polynucleotide comprises a polynucleotide encoding a polypeptide. Conveniently, the polynucleotide of this aspect of the invention and the polynucleotide expressing the polypeptide are operably linked. The polypeptide encoded may be any suitable polypeptide, but it is preferred if the polypeptide is one with therapeutic capabilities. The nature



of the therapeutic capability required depends on the circumstances, but examples may include nerve growth factors such as Nerve Growth Factor (NGF), Brain Derived Neurotrophic Factor (BDNF), polypeptides useful in treatment of neurodegenerative disease and the like. In an especially preferred embodiment, the polypeptide is the polypeptide whose amino acid  
5 sequence is given in Figure 1b or Figure 11, or a fragment or variant or fusion of said fragment or variant. A preferred variant of the polypeptide is one which has the sequence shown in Figure 18.

10 In an alternative preferred embodiment, the polypeptide is Protein Kinase D, or a mutant thereof such as are discussed below.

The polypeptide may also be any polypeptide which is detectable. Such polypeptides are useful in imaging cells which are able to use the  
15 polynucleotide whose nucleotide sequence is given in Figure 16 to promote transcription, or as markers by which to specifically target other reagents to the expressing cells.

In a further embodiment of this aspect of the invention, the polynucleotide  
20 is one able to replicate within a cell.

A fourth aspect of the present invention provides a host cell transformed with a polynucleotide vector construct of the present invention. The host cell can be either prokaryotic or eukaryotic. Bacterial cells are preferred  
25 prokaryotic host cells and typically are a strain of *E.coli* such as, for example, the *E.coli* strains DH5 available from Bethesda Research Laboratories Inc., Bethesda, MD, USA, and RR1 available from the American Type Culture Collection (ATCC) of Rockville, MD, USA (No. ATCC 31343). Preferred eukaryotic host cells include yeast, insect and

mammalian cells, preferably vertebrate cells such as those from a mouse, rat, monkey or human fibroblastic and kidney cell lines. Yeast host cells include YPH499, YPH500 and YPH501 which are generally available from Stratagene Cloning Systems, La Jolla, CA 92037, USA. Preferred  
5 mammalian host cells include Chinese hamster ovary (CHO) cells available from the ATCC as CRL 1658 and 293 cells which are human embryonic kidney cells. Preferred insect cells are Sf9 cells which can be transfected with baculovirus expression vectors.

10 Transformation of appropriate cell hosts with a DNA construct of the present invention is accomplished by well known methods that typically depend on the type of vector used. With regard to transformation of prokaryotic host cells, see, for example, Cohen *et al* (1972) *Proc. Natl. Acad. Sci. USA* 69, 2110 and Sambrook *et al* (1989) *Molecular Cloning, A*  
15 *Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY. Transformation of yeast cells is described in Sherman *et al* (1986) *Methods In Yeast Genetics, A Laboratory Manual*, Cold Spring Harbor, NY. The method of Beggs (1978) *Nature* 275, 104-109 is also useful. With regard to vertebrate cells, reagents useful in transfecting such cells, for  
20 example calcium phosphate and DEAE-dextran or liposome formulations, are available from Stratagene Cloning Systems, or Life Technologies Inc., Gaithersburg, MD 20877, USA.

Electroporation is also useful for transforming and/or transfecting cells and  
25 is well known in the art for transforming yeast cells, bacterial cells, insect cells and vertebrate cells.

For example, many bacterial species may be transformed by the methods described in Luchansky *et al* (1988) *Mol. Microbiol.* 2, 637-646

incorporated herein by reference. The greatest number of transformants is consistently recovered following electroporation of the DNA-cell mixture suspended in 2.5 PEB using 6250V per cm at 25  $\mu$ FD.

- 5 Methods for transformation of yeast by electroporation are disclosed in Becker & Guarente (1990) *Methods Enzymol.* **194**, 182.

Successfully transformed cells, ie cells that contain a DNA construct of the present invention, can be identified by well known techniques. For  
10 example, cells resulting from the introduction of an expression construct of the present invention can be grown to produce the polypeptide of the invention. Cells can be harvested and lysed and their DNA content examined for the presence of the DNA using a method such as that described by Southern (1975) *J. Mol. Biol.* **98**, 503 or Berent *et al* (1985)  
15 *Biotech.* **3**, 208. Alternatively, the presence of the protein in the supernatant can be detected using antibodies as described below.

In addition to directly assaying for the presence of recombinant DNA, successful transformation can be confirmed by well known immunological  
20 methods when the recombinant DNA is capable of directing the expression of the protein. For example, cells successfully transformed with an expression vector produce proteins displaying appropriate antigenicity.

Samples of cells suspected of being transformed are harvested and assayed  
25 for the protein using suitable antibodies.

The host cell may be a host cell within an animal body. Thus, transgenic animals which express a polypeptide of the fourth aspect of the invention by virtue of the presence of the transgene are included. Preferably, the

transgenic animal is a rodent such as a mouse. Transgenic animals can be made using methods well known in the art.

In one embodiment of the fourth aspect of the invention, the cells are transformed or transfected with a polynucleotide according to the first aspect of the invention or an expression vector according to the second aspect of the invention, and said cells further comprise a polynucleotide capable of expressing Protein Kinase D or a functional equivalent thereof. As discussed in more detail below, such cells may be useful for investigating the interaction of Protein Kinase D and Kidins220. Alternatively, cells capable of co-expressing Kidins220 and Protein Kinase D may be useful for investigating or producing Kidins220 in its phosphorylated form.

It will be appreciated that the Protein Kinase D expressed in said cells may be the form found in Nature, or may be a mutant form. Examples of useful mutant forms are constitutively active forms generated, for example, by a double mutation of residues serine 744 and serine 749 to glutamic acid, and a dominant negative form produced by mutation of aspartic acid 733 to alanine.

Functional equivalents include these mutants and other forms of PKD (such as variants and fusions) that are capable of phosphorylating Kidins220.

Other mutants of PKD which may be also useful in the invention are described in Iglesias *et al* (1998) *FEBS Lett.* 437:19-23; Iglesias and Rozengurt (1998) *J. Biol. Chem.* 273:410-416; Iglesias *et al* (1998) *J. Biol. Chem.* 273:27662-27667 all incorporated herein by reference and which refer to PKD cDNA and its sequence.

In addition to the transformed host cells themselves, the present invention also contemplates a culture of those cells, preferably a monoclonal (clonally homogenous) culture, or a culture derived from a monoclonal culture, in a  
5 nutrient medium.

A fifth aspect of the invention provide a polypeptide comprising the amino acid sequence given in Figure 1b or Figure 11 or a fragment or variant or fusion thereof, or a fusion of said fragment or variant provided that said  
10 polypeptide is not the polypeptide encoded by GenBank Accession No. AB033076.

Preferably, the polypeptide of the invention is not the polypeptide of WO 01/42285 identified as SEQ ID No 9.

15

Thus, a polypeptide of the invention includes a polypeptide whose amino acid sequence is given in Figure 1b or a polypeptide whose sequence is given in Figure 11. A polypeptide of the invention which is a variant of the amino acid sequence given in Figure 1b or Figure 11 includes that whose  
20 sequence is shown in Figure 18.

A polypeptide of the invention includes a polypeptide whose amino acid sequence is given in Figure 1b but wherein the "splicing cassette" (ie between residues 1140 and 1232 as indicated by the arrows in Figure 1b) is  
25 replaced by any one of the "cassettes" indicated in Figure 8 between the square brackets.

Preferably, a polypeptide of the invention is one which has at least 65% identity with the amino acid sequence given in Figure 1b or Figure 11 or

which sequence is as given in Figure 8. More preferably, the polypeptide has at least 70% or 80%, 85% or 90% identity, and even more preferably the polypeptide has at least 95% or 98% identity with the amino acid sequence shown in Figure 1b.

5

It will be appreciated that a polypeptide of the invention is not the polypeptide encoded by GenBank Accession No. AB033076, but a polypeptide of the invention may include the amino acid sequence encoded by GenBank Accession No. AB033076. Preferably, the polypeptide is not  
10 that of WO 01/42285 identified as SEQ ID No 9.

Therefore a polypeptide of the invention includes homologues of Kidins220 which are found in species other than rat, such as human, mouse, sheep and which are phosphorylation substrates of Protein Kinase D. It also includes  
15 variants of these that retain at least some activity of Kidins220.

As discussed above, the polypeptide of the invention includes human Kidins220.

20 Typically, the polypeptide of the invention is one which is able to interact with Protein Kinase D, and be phosphorylated on serine 919 (as numbered in Figure 1b) or serine 918 (as numbered in Figure 18). Preferably, the polypeptide of the invention is able to bind ATP.

25 It will be appreciated that the polypeptide of the invention may, or may not be phosphorylated. Where the polypeptide is phosphorylated, it may be on more than one residue. In particular, the polypeptide may, or may not be phosphorylated on serine 919 as numbered in Figure 1b or an equivalent

position. Both phosphorylated and non-phosphorylated forms of the polypeptide are included within the scope of the invention.

By "equivalent position" we mean an amino acid residue in a polypeptide, typically of the same type, which occupies a position which is the same as the given residue (in this case Ser 919) when the sequence of the polypeptide is optimally aligned, using the computer programs discussed above, with the reference sequence (in this case the Figure 1b sequence). Equivalent position, in this context, is well understood by the person skilled in the art. In the sequence of Figure 18, the equivalent serine is serine 918.

It will also be appreciated that the polypeptide of the invention may or may not have ATP or an analogue or functional equivalent thereof bound to it. Both ATP-bound and non ATP-bound forms are included within the scope of the invention.

"Fragments" of the Kidins220 polypeptide include polypeptides which comprise at least five consecutive amino acids of the polypeptide of the invention. Preferably, a fragment of the polypeptide comprises an amino acid sequence which is useful, for example, a fragment which retains activity of the polypeptide, or is useful as a peptide for producing an antibody which is specific for the Kidins220 polypeptide. Typically, the fragments have at least 8 consecutive amino acids, preferably at least 10, more preferably at least 12 or 15 or 20 or 30 or 40 or 50 consecutive amino acids of a polypeptide of the invention such as that shown in Figure 1b or the human amino acid sequence shown in Figure 11.

Particularly useful polypeptide fragments are those which comprise the Protein Kinase D phosphorylation recognition motif. The Protein Kinase D

phosphorylation recognition motif may be considered to include the motif (I/L)XRQMSJ, where X is any amino acid, J represents a hydrophobic residue such as isoleucine, leucine, phenylalanine, tryptophan, tyrosine or valine and (I/L) represents isoleucine or leucine. An example of such a polypeptide fragment includes the sequence RQMQRTITRQMSFDLTK.

It will be appreciated that the serine in the sequence RQMQRTITRQMSFDLTK (underlined; residue 919 as numbered in Figure 1b) which can be phosphorylated by Protein Kinase D, may be usefully changed to an amino acid residue other than serine. For example, serine 919 may be mutated to prevent phosphorylation of the polypeptide at this residue by Protein Kinase D. An example of such a mutation is RQMQRTITRQMAFDLTK. A conservative mutation of serine 919 may be to a threonine residue. Polypeptides which may mimic phosphorylated serine 919 are those in which serine 919 has been changed to an aspartate or glutamate residue. Polypeptides or fragments thereof containing such mutations are useful in competition assays, or as inhibitors of Protein Kinase D phosphorylating activity, or as epitopes for raising antibodies.

Typically, a useful fragment of the polypeptide of the invention is one which is antigenic. An example of an antigenic polypeptide fragment of the invention is CAASSESTGFGEERSIL. Such a fragment can be used to raise antibodies which are selective for the Kidins220 polypeptide, as described in more detail below.

In one embodiment of this aspect of the invention, the polypeptide fragment contains an amino acid sequence which is present in all of the splice variant forms of Kidins220. In other words, the fragment is a fragment of



Kidins220 outside of the splicing cassette shown in Figure 1b and therefore common to the different Kidins220 forms.

In another embodiment of this aspect of the invention, the polypeptide  
5 fragment contains an amino acid sequence which is present in the splicing  
cassette, ie the amino acid sequence between the arrows in Figure 1b or the  
amino acid sequences within the square brackets shown in Figure 8.  
Suitable fragments which are included in this embodiment are fragments  
including the amino acid sequence HPFYNRANIN (clone 1),  
10 HPFYNRAAVP (clone 2), DQNNGLAAVP (clone 3), RQVQKLQAAV  
(clone 4), CHRQLTVTEF (clone 5) and TEDAAEGLPS (clone 6).

It will be appreciated that fragments of the polypeptide of the invention are  
particularly useful when fused to other polypeptides, such as GST, Green  
15 Fluorescent Protein (GFP), Vesicular Stomatitis Virus Glycoprotein  
(VSVG) or Keyhole Limpet Haemocyanin (KLH). For example, a fusion of  
the fragment CAASSESTGFGEERSIL to GST may be used in methods  
well known in the art of antibody purification to purify a polyclonal  
antibody. Fusions of the polypeptide, or fusions of fragments or variants of  
20 the polypeptide of the invention are included in the scope of the invention.

Hence, a "fusion" of the Kidins220 polypeptide or a fragment or variant  
thereof provides a molecule comprising a polypeptide of the invention and a  
further portion. It is preferred that the said further portion confers a  
25 desirable feature on the said molecule; for example, the portion may be useful  
in detecting or isolating the molecule, or promoting cellular uptake of the  
molecule. The portion may be, for example, a biotin moiety, a radioactive  
moiety, a fluorescent moiety, for example a small fluorophore or a green  
fluorescent protein (GFP) fluorophore, as well known to those skilled in the

art. The moiety may be an immunogenic tag, for example a Myc tag, as known to those skilled in the art or may be a lipophilic molecule or polypeptide domain that is capable of promoting cellular uptake of the molecule or the interacting polypeptide, as known to those skilled in the art.

5

A sixth aspect of the invention provides a peptide comprising the sequence (I/L)XRQM(S/A)J wherein X is any amino acid, J is a hydrophobic amino acid such as isoleucine, leucine, phenylalanine, tryptophan, tyrosine or valine, (I/L) represents isoleucine or leucine and (S/A) represents serine or alanine and which is not the polypeptide encoded by the nucleotide  
10 sequence in GenBank Accession No AB033076 and which is not  $\mu$  peptide (AALVRQMSVAFFFK).

Preferably J is V or F.

15

When a serine is present rather than alanine at position (S/A), the peptide is believed to be a substrate for Protein Kinase D.

Preferably, the polypeptide comprises the amino acid sequence  
20 (I/L)XRQM(S/A)J and further amino acids to the N-terminus and C-terminus of this motif present in rat or human Kidins220 (see Figure 1b and Figure 11 and Figure 18). Typically, the peptide is at least 8 amino acids in length, preferably at least 10 amino acids but may be 15 or 20 or 25 or 30 or 35 or 40 amino acids in length.

25

Preferred peptides include the sequence RQMQRRTITRQMSFDLTK, or a minimum sequence thereof which can be phosphorylated by PKD may be useful as a substrate for Protein Kinase D since the serine residue in this sequence is phosphorylated in Kidins220 by Protein Kinase D.

Further preferred peptides include the sequence RQMQRITRQMAFDLTK, or a minimum sequence which can be phosphorylated by PKD where S has been changed to A, may be useful as a non-phosphorylatable control for Protein Kinase D, compared to the above polypeptides, since the phosphorylatable serine residue within the motif has been changed to alanine which cannot be phosphorylated.

A seventh aspect of the invention provides a peptide comprising the amino acid sequence CAASSESTGFGEERSIL or a variant or fragment or fusion thereof or a fusion of said fragment or variant or an analogue thereof provided that it is not the polypeptide encoded by the nucleotide sequence in GenBank Accession No AB033076.

Preferably the peptide is not a peptide of the polypeptide identified as SEQ ID No 9 in WO 01/42285.

The length preferences for these peptides is as discussed previously.

An eighth aspect of the invention provides a method of producing a peptide or polypeptide according to the fifth, sixth or seventh aspects of the invention comprising expressing a polynucleotide according to the second aspect of the invention.

In a preferred embodiment of the eighth aspect of the invention, the method of producing a peptide or polypeptide according to the fifth aspect of the invention comprises culturing a host cell according to the fourth aspect of the invention.

Methods of cultivating host cells and isolating recombinant proteins are well known in the art. It will be appreciated that, depending on the host cell, the Kidins220 produced may differ from that which can be isolated from nature. For example, certain host cells, such as yeast or bacterial  
5 cells, either do not have, or have different, post-translational modification systems which may result in the production of forms of Kidins220 which may be post-translationally modified in a different way to Kidins220 isolated from nature. In order to obtain Kidins220 which is post-translationally modified in a different way to human Kidins220 or rat  
10 Kidins220 it is preferred if the host cell is a non-human or non-rat host cell; more preferably it is not a mammalian cell.

It is preferred that the Kidins220 polypeptide is produced in a eukaryotic system, such as an insect cell.

15

According to a less preferred embodiment, the Kidins220 polypeptide can be produced *in vitro* using a commercially available *in vitro* translation system, such as rabbit reticulocyte lysate or wheatgerm lysate (available from Promega). Preferably, the translation system is rabbit reticulocyte  
20 lysate. Conveniently, the translation system may be coupled to a transcription system, such as the TNT transcription-translation system (Promega). This system has the advantage of producing suitable mRNA transcript from an encoding DNA polynucleotide in the same reaction as the translation. Conveniently, where the expressed polypeptide comprises one  
25 or more transmembrane domains, the translation system can be supplemented with a source of endoplasmic reticulum-derived membranes and folding chaperones, such as dog pancreatic microsomes, to allow synthesis of the polypeptide in a native conformation.

Preferably, the production method of the eighth aspect of the invention comprises a further step of isolating the Kidins220 produced from the host cell or from the *in vitro* translation mix. Preferably, the isolation employs an antibody which specifically binds the expressed polypeptide of the invention.

It will be understood that the invention comprises the Kidins220 polypeptide or the variant or fragment or fusion thereof, or a fusion of said variant or fragment obtainable by the methods herein disclosed.

10

A ninth aspect of the invention provides an antibody which specifically binds to a polypeptide whose amino acid sequence is given in Figure 1b or Figure 11, or a fragment or variant of the said polypeptide, for example, the polypeptide whose amino acid sequence is given in Figure 18. The antibodies include antibodies which are selective for Kidins220 and antibodies which are selective for the individual splice variant forms of Kidins220 as described. Thus, the antibodies of the invention include those that specifically bind to human Kidins220 or individual splice variant forms thereof.

20

Preferably, the antibody is not an antibody of WO 01/42285 which selectively binds the polypeptide whose sequence is identified as SEQ ID No 9 therein.

By "selective" we include antibodies which bind at least 10-fold more strongly to a polypeptide of the invention than to another polypeptide; preferably at least 50-fold more strongly and more preferably at least 100-fold more strongly. Such antibodies may be made by methods well known in the art using the information concerning the differences in amino acid

sequence between the amino acid sequence which is given in Figure 1b, or a variant thereof, and another polypeptide which is not a polypeptide of the invention. In particular, the antibodies may be polyclonal or monoclonal.

- 5    Suitable monoclonal antibodies which are reactive as said may be prepared by known techniques, for example those disclosed in *"Monoclonal Antibodies; A manual of techniques"*, H Zola (CRC Press, 1988) and in *"Monoclonal Hybridoma Antibodies: Techniques and Application"*, SGR Hurrell (CRC Press, 1982). Polyclonal antibodies may be produced which  
10   are polypepcific or monospecific. It is preferred that they are monospecific.

One embodiment provides an antibody which selectively binds the Kidins220 polypeptide whose amino acid is shown in Figure 1b or Figure 18 or a human equivalent thereof but does not bind a splice variant of  
15   Kidins220 as described.

A further embodiment provides an antibody which selectively binds a variant of Kidins220 as described but which antibody does not bind a polypeptide whose amino acid sequence is as given in Figure 1b or Figure  
20   18 or a human equivalent thereof.

A still further embodiment provides an antibody which selectively binds a particular variant or particular group of variants as described.

- 25   By "group of variants" is meant those variants which have some similarity in the way their amino acid sequence varies from the amino acid sequence given in Figure 1b. For example, the variant forms referred to as "Clone 4" and "Clone 5" in Figure 8 have the amino acid sequence PCDSGFNKQRQ inserted between residues [1178] and [1188] (as numbered in Figure 1b).

Hence, an antibody which selectively binds to the polypeptide sequence PCDSGFNKQRQ will react only with the group of variant forms referred to as "Clone 4" and "Clone 5" in Figure 8 and not with the amino acid sequence given in Figure 1, or with any of the other variants shown in Figure 8. Another example of a group of variants are the amino acid sequence given in Figure 1b and the variants referred to as Clone 3, 4, 5 and 6 in Figure 8, excluding the variants referred to as Clone 1 and 2. Both "Clone 1" and "Clone 2" in Figure 8 have a deletion of amino acids between residues [1140] and [1191]. Hence an antibody which selectively binds to the polypeptide sequence FFAPYLYTPR will not bind to the variant forms termed Clones 1 and 2 in Figure 8, but will bind to the sequence given in Figure 1 and to the variant forms referred to as Clones 3, 4, 5 and 6 in Figure 8.

Antibodies may distinguish between the different splice variants when able to bind to peptides unique to the different variants. Examples of such unique peptides are: clone 1, HPFYNRANIN; clone 2, HPFYNRRAAVP; clone 3, DQNNGLAAVP; clone 4, RQVQKLQAAV; clone 5, CHRQLTVTEF; clone 6, TEDAAEGLPS. For clones 1-3, these peptides represent the sequence spanning the point of fusion between the splice segments. The sequences suggested for clones 4-6 represent unique insertions in the variant. Such antibodies may be especially useful for monitoring a specific variant when used in conjunction with antibodies which recognise all forms of Kidins220 (such as one raised to the 21 amino acids at the C-terminus of Kidins220) so indicating the presence of the protein. As is discussed in more detail below, such antibodies may also be useful in medicine to monitor the presence or location of one or more forms of Kidins220, for example, in diagnosing cancer or monitoring the progression of a cancer or its treatment.

Hence, the use of an antibody according to the invention in medicine, for example, in diagnosis, prognosis or treatment of cancer, is contemplated.

- 5 The person skilled in the art can readily determine peptides which will be useful in making selective antibodies based on the information in Figures 1b and 8 and 18.

Further preferred antibodies are ones which distinguish Kidins220  
10 phosphorylated at serine 919 from rat Kidins220 (or an equivalent residue) not phosphorylated at serine 919. Such antibodies can be made, for example by immunising or selecting with suitable peptides which contain suitable portion of Kidin220 surrounding serine 919 and wherein the serine residue is phosphorylated and determining whether the antibody is selective  
15 for the phosphorylated form (ie does not substantially bound to the non-phosphorylated form).

Peptides in which one or more of the amino acid residues are chemically modified, before or after the peptide is synthesised, may be used providing  
20 that the function of the peptide, namely the production of specific antibodies *in vivo*, remains substantially unchanged. Such modifications included forming salts with acids or bases, especially physiologically acceptable organic or inorganic acids and bases, forming an ester or amid of a terminal carboxyl group, and attaching amino acid protecting groups such as N-t-  
25 butoxycarbonyl. Such modifications may protect the peptide from *in vivo* metabolism. The peptides may be present as single copies or as multiples, for example tandem repeats. Such tandem or multiple repeats may be sufficiently antigenic themselves to obviate the use of a carrier. It may be advantageous for the peptide to be formed as a loop, with the N-terminal



and C-terminal ends joined together, or to add one or more Cys residues to an end to increase antigenicity and/or to allow disulphide bonds to be formed. If the peptide is covalently linked to a carrier, preferably a polypeptide, then the arrangement is preferably such that the peptide of the invention forms a loop.

According to current immunological theories, a carrier function should be present in any immunogenic formulation in order to stimulate, or enhance stimulation of, the immune system. It is though that the best carriers embody (or, together with the antigen, create) a T-cell epitope. The peptides may be associated, for example by cross-linking, with a separate carrier, such as serum albumins, myoglobins, bacterial toxoids and keyhole limpit haemocyanin. More recently developed carriers which induce T-cell help in the immune response include the hepatitis-B core antigen (also called the nucleocapsid protein), presumed T-cell epitopes such as Thr-Ala-Ser-Gly-Val-Ala-Glu-Thr-Thr-Asn-Cys,  $\beta$ -galactosidase and the 163-171 peptide of interleukin-1. The latter compound may variously be regarded as a carrier or as an adjuvant or as both. Alternatively, several copies of the same or different peptides of the invention may be cross-linked to one another; in this situation there is no separate carrier as such, but a carrier function may be provided by such cross-linking. Suitably cross-linking agents include those listed as such in the Sigma and Pierce catalogues, for example glutaraldehyde, carbodiimide and succinimidyl 4-(N-maleimidomethyl)cyclohexane-1-carboxylate, the latter agent exploiting the -SH group on the C-terminal cysteine residue (if present).

If the peptide is prepared by expression of a suitable nucleotide sequence in a suitable host, then it may be advantageous to express the peptide as a

fusion product with a peptide sequence which acts as a carrier. Kabigen's "Ecosec" system is an example of such an arrangement.

Peptides may be synthesised by the Fmoc-polyamide mode of solid-phase peptide synthesis as disclosed by Lu *et al* (1981) *J. Org. Chem.* **46**, 3433 and references therein. Temporary N-amino group protection is afforded by the 9-fluorenylmethyloxycarbonyl (Fmoc) group. Repetitive cleavage of this highly base-labile protecting group is effected using 20% piperidine in N,N-dimethylformamide. Side-chain functionalities may be protected as their butyl ethers (in the case of serine threonine and tyrosine), butyl esters (in the case of glutamic acid and aspartic acid), butyloxycarbonyl derivative (in the case of lysine and histidine), trityl derivative (in the case of cysteine) and 4-methoxy-2,3,6-trimethylbenzenesulphonyl derivative (in the case of arginine). Where glutamine or asparagine are C-terminal residues, use is made of the 4,4'-dimethoxybenzhydryl group for protection of the side chain amido functionalities. The solid-phase support is based on a polydimethyl-acrylamide polymer constituted from the three monomers dimethylacrylamide (backbone-monomer), bisacryloylethylene diamine (cross linker) and acryloylsarcosine methyl ester (functionalising agent). The peptide-to-resin cleavable linked agent used is the acid-labile 4-hydroxymethyl-phenoxyacetic acid derivative. All amino acid derivatives are added as their preformed symmetrical anhydride derivatives with the exception of asparagine and glutamine, which are added using a reversed N,N-dicyclohexyl-carbodiimide/1-hydroxybenzotriazole mediated coupling procedure. All coupling and deprotection reactions are monitored using ninhydrin, trinitrobenzene sulphonic acid or isotin test procedures. Upon completion of synthesis, peptides are cleaved from the resin support with concomitant removal of side-chain protecting groups by treatment with 95% trifluoroacetic acid containing a 50% scavenger mix. Scavengers

commonly used are ethanedithiol, phenol, anisole and water, the exact choice depending on the constituent amino acids of the peptide being synthesised. Trifluoroacetic acid is removed by evaporation *in vacuo*, with subsequent trituration with diethyl ether affording the crude peptide. Any  
5 scavengers present are removed by a simple extraction procedure which on lyophilisation of the aqueous phase affords the crude peptide free of scavengers. Reagents for peptide synthesis are generally available from Calbiochem-Novabiochem (UK) Ltd, Nottingham NG7 2QJ, UK. Purification may be effected by any one, or a combination of, techniques  
10 such as size exclusion chromatography, ion-exchange chromatography and (principally) reverse-phase high performance liquid chromatography. Analysis of peptides may be carried out using thin layer chromatography, reverse-phase high performance liquid chromatography, amino-acid analysis after acid hydrolysis and by fast atom bombardment (FAB) mass  
15 spectrometric analysis.

The peptide of the invention may be linked to other antigens to provide a dual effect.

20 In a preferred embodiment of the invention, an antibody of the invention is capable of preventing or disrupting the interaction between Kidins220 polypeptide and Protein Kinase D.

It will be appreciated that other antibody-like molecules may be useful in  
25 the practice of the invention including, for example, antibody fragments or derivatives which retain their antigen-binding sites, synthetic antibody-like molecules such as single-chain Fv fragments (ScFv) and domain antibodies (dAbs), and other molecules with antibody-like antigen binding motifs.

Such antibody-like molecules are included by the term "antibody" as used above.

A further aspect of the invention provides a method of producing an antibody according to the ninth aspect of the invention.

Antibodies can be raised in an animal by immunising with an appropriate peptide. Appropriate peptides are described herein. Alternatively, with today's technology, it is possible to make antibodies as defined herein without the need to use animals. Such techniques include, for example, antibody phage display technology as is well known in the art. Appropriate peptides, as described herein, may be used to select antibodies produced in this way.

It will be appreciated that, with the advancements in antibody technology, it may not be necessary to immunise an animal in order to produce an antibody. Synthetic systems, such as phage display libraries, may be used. The use of such systems is included in the methods of the invention and the products of such systems are "antibodies" for the purposes of the invention.

It will be appreciated that such antibodies which recognise Kidins220 and variants or fragments thereof are useful research reagents. Suitably, the antibodies of the invention are detectably labelled, for example they may be labelled in such a way that they may be directly or indirectly detected. Conveniently, the antibodies are labelled with a radioactive moiety or a coloured moiety or a fluorescent moiety, or they may be linked to an enzyme. Typically, the enzyme is one which can convert a non-coloured (or non-fluorescent) substrate to a coloured (or fluorescent) product. The antibody may be labelled by biotin (or streptavidin) and then detected

indirectly using streptavidin (or biotin) which has been labelled with a radioactive moiety or a coloured moiety or a fluorescent moiety, or the like or they may be linked to any enzyme of the type described above.

- 5 A tenth aspect of the present invention provides a method of modulating the activity of the polypeptide whose amino acid sequence is given in Figure 1b or a fragment or variant or fusion thereof, the method comprising contacting the polypeptide with an antibody according to the seventh aspect of the invention. Thus, the activity of Kidins220 or a fragment or variant thereof  
10 can be modulated by contacting it with a suitable antibody as defined.

In a preferred embodiment, the antibody is specific for the form of Kidins220 which is non-phosphorylated on serine 919 (as numbered in Figure 1b). Addition of such an antibody may modulate the function of the  
15 polypeptide of the invention by inhibiting phosphorylation of the serine at residue 919 by preventing access of Protein Kinase D to the phosphorylation site. A lack of a phosphate group on serine 919 of Kidins220 may cause degradation of Kidins220 by an intracellular mechanism such as ubiquitination and proteosomal degradation.

20

In an alternative embodiment of the tenth aspect of the invention, the antibody may be specific for the form of Kidins220 which is not bound to ATP, or the form which is bound to ADP. Addition of such an antibody may modulate the function of the polypeptide of the invention by inhibiting  
25 the binding of ATP to the ATP-binding site. Such an inhibition may modulate the activity or function of the Kidins220 polypeptide by maintaining the polypeptide in a non-active state, or a constitutively active form, depending on the role of ATP in the function of Kidins220.

In a further embodiment of the tenth aspect of the invention, the activity of the Kidins220 polypeptide, or a variant or fragment or fusion thereof, or a fusion of said fragment or variant may be modified by binding of an antibody to a region involved in protein-protein interactions. Such a region of the Kidins220 protein is the proline-rich region which may be involved in binding adaptor modules such as SH3 and WW domains, or the C-terminal portion with a high serine and threonine content which may participate in interactions with serine/threonine binding modules. Further such regions are the ankyrin repeat regions located at amino acid 17 – 391 (as numbered in Figure 18).

An eleventh aspect of the invention comprises a method of modulating the function of the polypeptide whose amino acid sequence is given in Figure 1b or a fragment or variant or fusion thereof when in the presence of Protein Kinase D, the method comprising inhibiting the phosphorylating ability of Protein Kinase D. Typically, PKD is inhibited so that so that Kidins220 remains non-phosphorylated on serine 919 when in the presence of PKD. This inhibition of Protein Kinase D phosphorylating activity may be achieved by addition of an antibody which either specifically binds the active site of Protein Kinase D, or which specifically binds a site of activation or regulation, such as a phosphorylation or diacylglycerol binding domain (Iglesias and Rozengurt (1999) *FEBS Lett.* 454:53-56), of Protein Kinase D, which antibody binding prevents the activation of Protein Kinase D necessary for its Kidins220-phosphorylating activity. A suitable target phosphorylation site on Protein Kinase D includes the two activation loop sites in the carboxy-terminal catalytic domain described by Iglesias *et al* (1998) *J. Biol. Chem.* 273:27662-27667) and Waldron *et al* (1999). Another site which may be a suitable target of antibody binding in order to inhibit the phosphorylating activity of Protein Kinase D is the amino-

terminal pleckstrin homology (PH) domain of the protein that regulates its activity (Iglesias *et al* (1998) *J. Biol. Chem* 273:410-416).

5 A twelfth aspect of the invention provides a method of modulating the activity of the polypeptide whose amino acid sequence is given in Figure 1b or a fragment or variant or fusion thereof, the method comprising reducing the ability of said polypeptide to bind ATP.

10 In a preferred embodiment, the ability of the Kidins220 polypeptide to bind to ATP is modulated by addition of a small molecule which binds the ATP-binding 'pocket', or a non-hydrolysable ATP analogue. Such an analogue is 8-azido-ATP, which in the presence of UV light covalently modifies ATP-binding proteins and prevents binding of a hydrolysable ATP molecule.

15 A thirteenth aspect of the invention provides a method of identifying an agent which modulates the promoter activity of the polynucleotide comprising the nucleotide sequence given in Figure 16 or a fragment or variant thereof, said method comprising i) introducing said polynucleotide into a suitable cell; ii) exposing a test agent to said cell and; iii) detecting  
20 whether (and, optionally, to what extent) the activity of the polynucleotide has been altered.

From the sequence presented in Figure 16, the skilled person can identify  
25 the promoter sequence or regulatory region contained therein using no more than routine experimentation.

Typically, the polynucleotide further comprises a polynucleotide capable of expressing a polypeptide, which polypeptide is a detectable one. Suitable

detectable polypeptides include green fluorescent protein, enzymes capable of producing coloured or detectable products from a colourless or undetectable substrate (for example, horseradish peroxidase), or polypeptide for which there is a specific antibody readily available. Agents which  
5 produce an increase in the amount of the polypeptide expressed are stimulators of the promoter activity of the polynucleotide. Agents which produce a decrease in the amount of polypeptide expressed are inhibitors of the promoter activity of the polypeptide.

10 Suitable cells are typically neuronal cells, such as PC12 cells or other neuronal cell lines, but may also be primary neuronal cells.

Agents identified by this method may be useful in enhancing (in the case of stimulators of promoter activity) or decreasing (in the case of inhibitors of  
15 promoter activity) the level of expression of polypeptides in neuronal cells. Modulation of expression of a polypeptide may be useful in the treatment of neurogenerative disease where the disease arises as a result of excessive or inadequate levels of a polypeptide within the neuronal cells.

20 Agents identified by this method may also be useful in manipulating the level of expression of polypeptides in neuronal cells in culture.

A fourteenth aspect of the invention provides a method of identifying an agent which modulates the activity of the polypeptide whose amino acid  
25 sequence is given in Figure 1b, or a fragment or variant or fusion thereof, or a fusion of said fragment or variant said method comprising: i) exposing a test agent to said polypeptide and; ii) detecting whether (and, optionally, to what extent) the activity of said polypeptide has been altered.



Typically, the polypeptide is the polypeptide whose amino acid sequence is given in figure 1b or a variant or fusion or fragment which is a functional equivalent thereof.

- 5 By "functional equivalent" we mean that the said component is one whose activity is substantially the same as the Kidins220 polypeptide activity which is being sought to be modified by the method, but which may be more convenient to use in an assay.
- 10 By "determining whether (and, optionally, to what extent) said polypeptide has been altered" we mean that the activity of the polypeptide when exposed to the test substance is compared to the activity of the polypeptide when not exposed to that test substance under the same assay conditions. It is understood that the activity may be increased, or decreased as a result of
- 15 exposure to the test substance or it may not be altered.

Typically, the activity of said Kidins220 polypeptide which is measured in the assay may be any one or more of its ability to interact with another polypeptide, its ability to become ubiquitinated, its ability to become

20 phosphorylated, its ability to bind ATP and its ability to dissociate from ADP. An example of another polypeptide with which Kidins220 may interact is Protein Kinase D. Kidins220 may interact with TrkA and p75 neurotrophin receptors. An example of an amino acid residue in Kidins220 which may become phosphorylated is serine 919 as numbered in Figure 1b.

25

Phosphorylation and ubiquitination can be measured easily using a standard ELISA assay with antibody against the phosphorylated form of Kidins220, or against ubiquitin. Protein interaction may be measured using standard

immunoprecipitation techniques where an antibody used in the immunoprecipitation is detectably labelled, or becomes detectably labelled.

The enhancement or disruption or prevention of an interaction between the  
5 Kidins220 polypeptide and another polypeptide can be measured *in vitro*  
using methods well known in the art of biochemistry and include any  
methods which can be used to assess protein-protein interactions. Suitable  
methods include co-purification, ELISA and co-immunoprecipitation  
10 methods. Further suitable methods may include Scintillation Proximity  
Assays, as well as other techniques well known to those skilled in the art.

The said interaction can also be measured within a cell, for example using  
the yeast two hybrid system or green fluorescent protein (GFP)-based  
systems as is well known in the art.

15 In one embodiment, a host cell according to the fourth aspect of the  
invention which comprises a polynucleotide capable of expressing a  
Kidins220 polypeptide according to the fifth aspect of the invention, and a  
further polynucleotide capable of expressing Protein Kinase D may be  
20 suitable for screening for agents which enhance or disrupt or prevent an  
interaction between Kidins220 and Protein Kinase D.

Such agents identified by said screening method may be useful in research  
into the effect of Protein Kinase D phosphorylation in cellular function.  
25 Such agents may also be useful as medicaments for treating pathologies  
which arise due to aberrant phosphorylation by Protein Kinase D.

Agents identified by the screening method may be useful in their own right, or may be useful as lead compounds for the development of other compounds which display greater efficacy.

- 5 A fifteenth aspect of the invention provides a method of identifying an agent able to modulate the activity of Protein Kinase D, the method comprising: i) exposing Protein Kinase D or a functional equivalent thereof to a peptide or polypeptide comprising the amino acid sequence (I/L)XRQMSJ, where X is any amino acid, J represents a hydrophobic  
10 residue such as isoleucine, leucine, phenylalanine, tryptophan, tyrosine or valine and (I/L) represents isoleucine or leucine; and ii) detecting in the presence or absence of the test agent whether (and, optionally, to what extent) said peptide or polypeptide has been phosphorylated, provided that the peptide is not  $\mu$  peptide (AALVRQMSVAFFFK).

15

- Additionally or alternatively, the polypeptide exposed to PKD may comprise a polypeptide comprising at least five consecutive residues including residue 919 of the polypeptide as numbered in Figure 1b. Preferably, the peptide is one including the sequence  
20 RQMQRITRQMSFDLTK or a minimum sequence thereof which can be phosphorylated by PKD.

Typically, the test agent of step i) is exposed to the Kidins220 polypeptide, or variant or fragment or fusion or fusion of said fragment or variant, prior  
25 to being exposed to Protein Kinase D.

In one embodiment, the test agent is exposed to Protein Kinase D prior to being exposed to the Kidins220 polypeptide, or variant or fragment or fusion or fusion of said fragment or variant.

Functional equivalents of Protein Kinase D include forms of PKD such as variants and fusions, as discussed above, which are capable of phosphorylating Kidins220. It will be appreciated that the presence of the  
5 test agent may affect the activity of Protein Kinase D either positively or negatively. By this we mean that the ability of Protein Kinase D to phosphorylate the target polypeptide may be enhanced, or it may be decreased. Where the phosphorylating activity of Protein Kinase D is enhanced in the presence of the test agent, the level of phosphorylation of  
10 the polypeptide will be greater than that achieved in the absence of the test agent.

Detection of the presence and/or amount of phosphate associated with the Kidins220 polypeptide after the assay may be conducted by any convenient  
15 means. Generally detection may be accomplished by using labelled (e.g. radiolabelled) phosphate in free solution or attached to a substrate, and comparing the amount associated with the substrate before and after the assay. Alternatively, antibodies which distinguish the phosphorylated form of Kidins220 and the non-phosphorylated form can be used.

20

It will be appreciated that the invention provides screening assays for drugs which may be useful in modulating the activity of Kidins220 or its interactions, or the activity of Protein Kinase D. The compounds identified in the methods may themselves be useful as a drug or they may represent  
25 lead compounds for the design and synthesis of more efficacious compounds.

A sixteenth aspect of the invention provides a compound identifiable or identified by the screening methods of the invention.

The invention provides the use of Kidins220 (*in vitro* or *in vivo*), or a fragment or variant or fusion thereof, or a fusion of said variant or fragment, or a functional equivalent thereof in the modulation of Protein Kinase D  
5 function. A composition comprising such agents is also included in the present invention.

Further aspects of the invention provide agents according to the sixteenth aspect of the invention, polynucleotides according to the first and second  
10 aspects of the invention and polypeptides or peptides according to the fourth, fifth, sixth or seventh aspects of the invention for use in medicine.

Preferable, the use in medicine is not a use in inflammation, trauma or cell proliferation.

15

Peptides based on Kidins220 as described herein which contain a phosphorylation site for PKD may be used to compete for PKD-mediated phosphorylation of Kidins220. These peptides may be used to down-regulate the activity of PKD on Kidins220 and other substrates of PKD.

20

Said agents or polynucleotides or polypeptides are useful for treatment of neurodegenerative disease. Neurodegenerative disease may arise due to a patient possessing an aberrant form of either PKD or Kidins220. Impairment in the function of either PKD or Kidins220 might determine  
25 failure in neuronal differentiation or failure in neuronal maintenance. Administration of an agent able to increase function of Kidins220, or of a polynucleotide or polypeptide able to restore normal Kidins220 function or expression level within the patient may be an effective treatment of the neurodegenerative disease.

In one embodiment, treatment of the neurodegenerative disease may be by gene therapy. A preferred expression vector for use in such gene therapy is one suitable for use in gene therapy according to the second aspect of the invention. Where it is known that the neurodegenerative disorder has arisen through inadequate function of PKD, an expression vector suitable for use in treatment of the disease may be one which expresses a fully functional form of PKD or a suitable equivalent thereof (for example, a fragment of the full length PKD polypeptide which is able to phosphorylate Kidins220) which is able to restore the missing PKD function to the patient.

Gene therapy may be carried out according to generally accepted methods, for example, as described by Friedman, 1991. A virus or plasmid vector (see further details below), containing a copy of the gene linked to expression control elements and capable of replicating inside neuronal cells, is prepared. Suitable vectors are known, such as disclosed in US Patent 5,252,479 and PCT published application WO 93/07282. The vector is then injected into the patient, either locally at the site of the neurodegeneration, or systemically. If the transfected PKD or Kidins220 polynucleotide is not permanently incorporated into the genome of each of the targeted neuronal cells, the treatment may have to be repeated periodically.

Gene transfer systems known in the art may be useful in the practice of the gene therapy methods of the present invention. These include viral and nonviral transfer methods. A number of viruses have been used as gene transfer vectors, including papovaviruses, eg SV40 (Madzak *et al*, 1992), adenovirus (Berkner, 1992; Berkner *et al*, 1988; Gorziglia and Kapikian, 1992; Quantin *et al*, 1992; Rosenfeld *et al*, 1992; Wilkinson *et al*, 1992; Stratford-Perricaudet *et al*, 1990), vaccinia virus (Moss, 1992), adeno-

associated virus (Muzyczka, 1992; Ohi *et al*, 1990), herpesviruses including HSV and EBV (Margolskee, 1992; Johnson *et al*, 1992; Fink *et al*, 1992; Breakfield and Geller, 1987; Freese *et al*, 1990), and retroviruses of avian (Brandyopadhyay and Temin, 1984; Petropoulos *et al*, 1992), murine  
5 (Miller, 1992; Miller *et al*, 1985; Sorge *et al*, 1984; Mann and Baltimore, 1985; Miller *et al*, 1988), and human origin (Shimada *et al*, 1991; Helseth *et al*, 1990; Page *et al*, 1990; Buchschacher and Panganiban, 1992). To date most human gene therapy protocols have been based on disabled murine retroviruses.

10

Nonviral gene transfer methods known in the art include chemical techniques such as calcium phosphate coprecipitation (Graham and van der Eb, 1973; Pellicer *et al*, 1980); mechanical techniques, for example microinjection (Anderson *et al*, 1980; Gordon *et al*, 1980; Brinster *et al*,  
15 1981; Constantini and Lacy, 1981); membrane fusion-mediated transfer *via* liposomes (Felgner *et al*, 1987; Wang and Huang, 1989; Kaneda *et al*, 1989; Stewart *et al*, 1992; Nabel *et al*, 1990; Lim *et al*, 1992); and direct DNA uptake and receptor-mediated DNA transfer (Wolff *et al*, 1990; Wu *et al*, 1991; Zenke *et al*, 1990; Wu *et al*, 1989b; Wolff *et al*, 1991; Wagner *et al*,  
20 1990; Wagner *et al*, 1991; Cotten *et al*, 1990; Curiel *et al*, 1991a; Curiel *et al*, 1991b).

Other suitable systems include the retroviral-adenoviral hybrid system described by Feng *et al* (1997) *Nature Biotechnology* **15**, 866-870, or viral  
25 systems with targeting ligands such as suitable single chain Fv fragments.

In an approach which combines biological and physical gene transfer methods, plasmid DNA of any size is combined with a polylysine-conjugated antibody specific to the adenovirus hexon protein, and the

resulting complex is bound to an adenovirus vector. The trimolecular complex is then used to infect cells. The adenovirus vector permits efficient binding, internalization, and degradation of the endosome before the coupled DNA is damaged.

5

Liposome/DNA complexes have been shown to be capable of mediating direct *in vivo* gene transfer. While in standard liposome preparations the gene transfer process is nonspecific, localized *in vivo* uptake and expression have been reported in tumour deposits, for example, following direct *in situ* administration (Nabel, 1992).

10

Gene transfer techniques which target DNA directly to tissues, eg epithelial cells of the breast or ovaries, is preferred. Receptor-mediated gene transfer, for example, is accomplished by the conjugation of DNA (usually in the form of covalently closed supercoiled plasmid) to a protein ligand *via* polylysine. Ligands are chosen on the basis of the presence of the corresponding ligand receptors on the cell surface of the target cell/tissue type. One appropriate receptor/ligand pair for introduction of the therapeutic gene into breast tumour cells may include the estrogen receptor and its ligand, estrogen (and estrogen analogues). These ligand-DNA conjugates can be injected directly into the blood if desired and are directed to the target tissue where receptor binding and internalization of the DNA-protein complex occurs. To overcome the problem of intracellular destruction of DNA, coinfection with adenovirus can be included to disrupt endosome function.

15

20

25

A further embodiment of the invention provides a method of treating a neurodegenerative disease, such as failure of adrenal function, comprising the step of administering to the patient an effective amount of either



Kidins220 or PKD polypeptide or a fragment or variant or fusion thereof to ameliorate the neurodegenerative disease.

Further aspects of the invention provide a pharmaceutical composition  
5 comprising an agent identified or identifiable by the methods of the  
sixteenth aspect of the invention or a polynucleotide according to the first or  
second aspects of the invention or a polypeptide according to the fifth, sixth  
or seventh aspects of the invention and a pharmaceutically acceptable  
carrier.

10

By "pharmaceutically acceptable" is included that the formulation is sterile  
and pyrogen free. Suitable pharmaceutical carriers are well known in the art  
of pharmacy.

15 Further aspects of the invention provide kits of parts that are useful in  
carrying out the screening methods of the invention or for phosphorylating a  
polypeptide of the invention by Protein Kinase D. Such a kit may be useful  
in research where the effect of phosphorylation of serine 919 in Kidins220  
is of interest in investigating the cell biology of a disease state.

20

Thus, one kit includes Protein Kinase D or a functional equivalent thereof  
and a polypeptide whose amino acid sequence is known in Figure 1b or a  
fragment or variant or fusion thereof. A preferred variant is the polypeptide  
whose amino acid sequence is shown in Figure 18.

25

Preferably, the Kidins220-related polypeptide present in the kit comprises  
an amino acid sequence of at least five consecutive amino acids including  
serine 919 as numbered in Figure 1b. Typically, however, the kit contains a  
polypeptide which contains the amino acid sequence

RQMQRITRQMSFDLTK. The kit may additionally or alternatively contain a polypeptide comprising the amino acid sequence (I/L)XRQMSJ, where X is any amino acid, J represents a hydrophobic residue such as isoleucine, leucine, phenylalanine, tryptophan, tyrosine or valine and (I/L)  
5 represents isoleucine or leucine.

The Protein Kinase D of the kit may be a constitutively active form of Protein Kinase D, which may be useful in research into the effects of aberrant phosphorylation of Kidins220. A constitutively active form may  
10 be made by a double mutation of Protein Kinase D as described above.

The kit may, or may not further comprise a means of detecting phosphorylation of a protein. Such a means may be by provision of a labelled (eg radiolabelled) phosphate compound such as ATP, suitable as a  
15 phosphate donor in a phosphorylation reaction, or an antibody which specifically binds the phosphorylated form of the Kidins220 polypeptide and does not bind the non-phosphorylated polypeptide. Conveniently, the antibody would be labelled with a detectable group as described above. Conveniently, the kit comprises an appropriate means for detecting the  
20 label, such as autoradiograph film or an enzyme-linked chemiluminescent assay. Some of the kit components may be localised on a surface, such as a blotting membrane, or an assay plate for ELISA etc, although the assay can be carried out in solution.

25 A still further aspect of the invention provides a mutant animal wherein if the organism contains in its genome a polynucleotide which encodes a polypeptide comprising the amino acid sequence in Figure 1b or a fragment or a variant thereof the polynucleotide in the mutant organism is mutated or absent.

Such mutant animal can be made by gene knockout and gene replacement technology as is well known in the art. The organism may be any suitable organism such as *Drosophila* or a rodent. Preferably, however, the  
5 organism is a non-human mammal such as a rodent such as a rat or mouse. A suitable method of making such mutants in *C.elegans* is given in Kraemer *et al* (1999), *Curr. Biol.* 9:1009-1018, and in mice in Torres and Kuehn, (1997) *Laboratory Protocols for Conditional Gene Targeting*, Oxford University Press. The method of RNA interference for gene silencing is  
10 described in Bosher and Labouesse (2000) *Nature Cell Biol.* 2:E31-E36.

A yet further aspect of the invention provides a method of diagnosing cancer in a patient comprising the steps of

- (i) determining the level of expression of the Kidins220 gene in a  
15 sample from the patient containing protein or nucleic acid;
- (ii) comparing the level of expression of step (i) with that in a control sample.

The level of expression of the Kidins220 gene in the sample may be  
20 determined by any convenient means. For example, the amount of mRNA encoding the Kidins220 polypeptide can be detected using detectably labelled polynucleotide probes which are capable of selectively hybridising to the mRNA. Suitable polynucleotide probes are those according to the present invention which are described in more detail above. Alternatively  
25 or additionally, the amount of the Kidins220 polypeptide may be detected using molecules which are capable of selectively binding to the Kidins220 polypeptide. Suitable molecules include antibodies selective for Kidins220 of the present invention, such as those described in more detail above.

According to one embodiment of this aspect of the invention, the subcellular location of the Kidins220 gene product is also determined in addition to the level of expression.

- 5 A change in the cellular location of the Kidins220 gene product (the Kidins220 polypeptide) from the plasma membrane to the nucleus or cytoplasm may indicate that the cell is a cancerous or pre-cancerous one. The change may be observed as an increase in the proportion of nuclear or cytoplasmic Kidins220 polypeptide compared to a control sample, or as a  
10 presence of nuclear or cytoplasmic Kidins220 compared to an absence of nuclear or cytoplasmic Kidins220 in the control sample.

By "cellular location" we include the meaning of which cellular compartment the polypeptide is observed to be in. The term "cellular  
15 compartment" is well understood in the art, and includes the plasma membrane, the cytosol, the nucleus, endoplasmic reticulum, Golgi complex, peroxisome and so on. The presence of nuclear or cytoplasmic Kidins220 polypeptide, or an increase in the proportion of Kidins220 polypeptide in the nucleus or cytoplasm, may, in addition to an increase in the expression  
20 level of the Kidins220 gene, further indicate that the cancer is progressing or developing into a more advanced cancerous stage. In this embodiment, it may be useful to compare the cellular location of the Kidins220 polypeptide to that of a control sample.

- 25 Methods of determining cellular location for a polypeptide are known in the art of cell biology, and include immunofluorescence and confocal microscopy, and subcellular fractionation. Methods for quantitating the (relative) amounts of polypeptide in specific subcellular locations are also known, and include the use of computer software packages which analyse

and quantitate the data produced by, for example, labelled antibodies in immunofluorescence.

According to a further embodiment, the phosphorylation status of the  
5 Kidins220 polypeptide in the sample from the patient is also determined in addition to the level of expression and compared to the phosphorylation status in a control sample. As discussed above and in Kong *et al* (2001) *J. Neurosci.* 21:176-185, the Kidins220 polypeptide may be phosphorylated on several residues, including on serine 919 (or its equivalent at serine 918  
10 shown in Figure 18) and on at least one tyrosine. A change in the phosphorylation status of the Kidins220 polypeptide in a sample from the patient (for example, an increase or a decrease in the total number of phosphorylated residues or a change in the phosphorylation of any one or more residues) compared to the phosphorylation status of Kidins220 in a  
15 control sample may be indicative of a cancerous or pre-cancerous state.

By "phosphorylation status is determined" we include the meaning that the number of phosphorylated residues in Kidins220 is determined, and/or the type of residues (for example, tyrosine or serine or threonine) which are  
20 phosphorylated and/or the presence or absence of a phosphate group on a specific residue is determined.

Methods for detecting phosphorylation on amino acid residues are well known in the art, and include the use of anti-phosphopeptide antibodies, for example, the  
25 anti-phosphotyrosine antibody pY99 available from Santa Cruz Biotechnology, Santa Cruz, CA, or anti-phosphoserine antibodies which selectively bind phosphorylated serine 919 (or its equivalent) as described above.

Glioblastomas are a less differentiated type of astrocytic tumour, and originate from astrocytomas and anaplastic astrocytomas undergoing neoplastic progression. The lack of differentiation means a particularly unfavourable prognosis and makes histopathological classification and grading particularly difficult. Routine treatment for glioblastoma multiforme involves surgical tumour reduction and cycles of radiation and chemotherapeutic treatment; however, these are mostly ineffective. The Kidins220 polypeptide is over-expressed in the cells of a variety of tumours, especially in glioblastoma multiforme and prostate cancers. Hence, detecting over-expression of the Kidins220 polypeptide, or detecting over-expression of the encoding polynucleotide may be indicative of a cancerous or precancerous state in those cells and allow diagnosis, or aid in the diagnosis, of cancer, especially prostate cancers and cancers of astrocytic origin such as glioblastoma multiforme. Early, and correct diagnosis of glioblastoma multiforme in particular offers the possibility for a prompt and suitable treatment which may produce a more favourable outcome.

The sample may be any suitable sample which includes protein or nucleic acid as required. The sample may be directly derived from the patient, for example, by biopsy of a tissue which may be associated with aberrant vascular development, or it may be derived from the patient from a site remote from the tissue, for example because cells from the tissue have migrated from the tissue to other parts of the body (eg tumour tissue may be taken from the primary tumour or from metastases). Alternatively, the sample may be indirectly derived from the patient in the sense that, for example, the tissue or cells therefrom may be cultivated *in vitro*, or cultivated in a xenograft model; or the nucleic acid sample may be one which has been replicated (whether *in vitro* or *in vivo*) from nucleic acid from the original source from the patient. Thus, although the nucleic acid

derived from the patient may have been physically within the patient, it may alternatively have been copied from nucleic acid which was physically within the patient.

- 5 Where the cellular location of the Kidins220 gene product is also determined, the sample from the patient is preferably one which contains cells.

Where the phosphorylation status is determined, the sample from the patient is preferably one which contains protein.

10

The term "over-expression" is well understood in the art, and we include the meaning that the level of expression is detectably higher than in a control sample, as defined below.

- 15 By "control sample" we include a sample from a tissue which is not believed to be cancerous. The tissue may, or may not, be from the same individual who is the patient being diagnosed. Where the tissue is from the same individual who is being diagnosed it is preferred if the tissue is not one in which cancer has been suspected or found.

20

A further aspect of the invention provides a method of monitoring the progression of cancer in a patient comprising the steps of

- (i) determining the level of expression of the Kidins220 gene in a sample from the patient containing protein or nucleic acid;
- 25 (ii) comparing the level of expression of step (i) with that in a previous sample from the same patient.

By "Kidins220 gene" we mean a gene which encodes a polypeptide of the invention. For example the Kidins220 gene is one which encodes the

polypeptides whose sequence is shown in Figures 1b, 11 and 18 and variants thereof.

Over-expression of the Kidins220 gene may be indicative of a cancerous or  
5 pre-cancerous state. Hence, an increase in the expression level (such as over-expression or greater degree of over-expression) may indicate that the tumour is growing or developing into a more advanced cancerous stage.

As discussed above in relation to the method for diagnosing cancer, the  
10 level of gene expression in the sample may be determined by any convenient technique known in the art.

According to one embodiment of this aspect, the cellular location of the Kidins220 gene product (Kidins220 polypeptide) is also determined in  
15 addition to the level of expression and compared to a previous sample.

Methods of determining the location and amount of polypeptide are discussed above.

20 According to another embodiment of this aspect, the phosphorylation status of the Kidins220 polypeptide is determined in addition to the level of expression, and compared to a previous sample from the patient. Methods of determining the phosphorylation status of Kidins220 are discussed above. A change in the phosphorylation status (as defined above) may indicate a  
25 progression or development in the cancer.

Hence, the invention provides a use for an antibody capable of selectively binding to the Kidins220 polypeptide in medicine, and especially in the manufacture of a diagnostic reagent for diagnosing or prognosing cancer.



As is discussed above, Kidins220 has several splice variants. Hence, the choice of complementary polynucleotide or antibody which is to be used in determining the level of Kidins220 polynucleotide or polypeptide in a sample may depend on the particular splice variant which it is desired to detect, or may be designed such that it does not discriminate between Kidins220 splice variants.

Although as discussed above determination of phosphorylation status or determination of cellular location may be done in combination with the determination of the level of expression of the Kidins220 gene, they may suitably be done without determining the level of expression.

Thus, a further aspect of the invention provides a method of diagnosing cancer in a patient comprising the steps of

- (i) determining the phosphorylation status of the Kidins220 polypeptide in a sample from the patient containing protein;
- (ii) comparing the phosphorylation status of step (i) with that in a control sample.

A yet further aspect of the invention provides a method of diagnosing cancer in a patient comprising the steps of

- (i) determining the cellular location of the Kidins220 polypeptide in a sample from the patient containing cells;
- (ii) comparing the cellular location of step (i) with that in a control sample.

As discussed above, the cellular location of the Kidins220 polypeptide may change in a cancerous or pre-cancerous cell, and may move from the plasma membrane to the cytoplasm and/or the nucleus.

- 5 A still further aspect of the invention provides a method of monitoring the progression of cancer in a patient comprising the steps of
- (i) determining the phosphorylation status of the Kidins220 polypeptide in a sample from the patient containing protein;
  - (ii) comparing the phosphorylation status of step (i) with that in a
- 10 control sample.

A still further aspect of the invention provides a method of monitoring the progression of cancer in a patient comprising the steps of

- (i) determining the cellular location of the Kidins220 polypeptide
- 15 in a sample from the patient containing cells;
- (ii) comparing the cellular location of step (i) with that in a
- control sample.

As discussed above, the cellular location of the Kidins220 polypeptide may

20 change in a cancerous cell, and the polypeptide may move from the plasma membrane to the cytoplasm and/or the nucleus, or may no longer be targeted to the plasma membrane during expression of the encoding polynucleotide and instead be directed to the cytoplasm and/or the nucleus. Such a change may be indicative of a change in the development stage of

25 the cancer, and thereby provide an indication of the degree success or failure of a treatment regime being administered to the patient following an earlier diagnosis of cancer.

Methods of determining cellular location and the phosphorylation status of Kidins220 are discussed above.

Suitable samples and control sample are as described above.

5

It will be appreciated that in the above methods of diagnosis of cancer or of monitoring the progression of cancer, it may be useful to determine the cellular location or phosphorylation status of a specific variant of Kidins220. Such variants are described above. Preferred variants have the  
10 sequence shown in Figure 18 and/or any of the splice forms shown in Figure 8.

Another aspect of the invention provides a method of treating cancer in a patient comprising administering to the patient an agent which inhibits  
15 function of the Kidins220 gene or its product.

The agent which inhibits the function of the Kidins220 gene may be any convenient agent. The agent may inhibit the function of the Kidins220 gene at one or more of several levels. For example, it may inhibit transcription of  
20 the gene or translation of the mRNA or it may inhibit the function of the protein product. An example of a suitable agent includes a polynucleotide which is antisense to a polynucleotide according to the first or third aspect of the invention.

25 Conveniently, the polynucleotide is antisense to a polynucleotide encoding a polypeptide according to the fifth aspect of the invention.

Antisense oligonucleotides are single-stranded nucleic acids, which can specifically bind to a complementary nucleic acid sequence. By binding to

the appropriate target sequence, an RNA-RNA, a DNA-DNA, or RNA-DNA duplex is formed. These nucleic acids are often termed "antisense" because they are complementary to the sense or coding strand of the gene. Recently, formation of a triple helix has proven possible where the  
5 oligonucleotide is bound to a DNA duplex. It was found that oligonucleotides could recognise sequences in the major groove of the DNA double helix. A triple helix was formed thereby. This suggests that it is possible to synthesise a sequence-specific molecules which specifically bind double-stranded DNA *via* recognition of major groove hydrogen binding  
10 sites.

By binding to the target nucleic acid, the above oligonucleotides can inhibit the function of the target nucleic acid. This could, for example, be a result of blocking the transcription, processing, poly(A)addition, replication,  
15 translation, or promoting inhibitory mechanisms of the cells, such as promoting RNA degradations.

Antisense oligonucleotides are prepared in the laboratory and then introduced into cells, for example by microinjection or uptake from the cell  
20 culture medium into the cells, or they are expressed in cells after transfection with plasmids or retroviruses or other vectors carrying an antisense gene. Antisense oligonucleotides were first discovered to inhibit viral replication or expression in cell culture for Rous sarcoma virus, vesicular stomatitis virus, herpes simplex virus type 1, simian virus and  
25 influenza virus. Since then, inhibition of mRNA translation by antisense oligonucleotides has been studied extensively in cell-free systems including rabbit reticulocyte lysates and wheat germ extracts. Inhibition of viral function by antisense oligonucleotides has been demonstrated *in vitro* using oligonucleotides which were complementary to the AIDS HIV retrovirus

RNA (Goodchild, J. 1988 "Inhibition of Human Immunodeficiency Virus Replication by Antisense Oligodeoxynucleotides", *Proc. Natl. Acad. Sci. (USA)* 85(15), 5507-11). The Goodchild study showed that oligonucleotides that were most effective were complementary to the poly(A) signal; also effective were those targeted at the 5' end of the RNA, particularly the cap and 5' untranslated region, next to the primer binding site and at the primer binding site. The cap, 5' untranslated region, and poly(A) signal lie within the sequence repeated at the ends of retrovirus RNA (R region) and the oligonucleotides complementary to these may bind twice to the RNA.

Typically, antisense oligonucleotides are 15 to 35 bases in length. For example, 20-mer oligonucleotides have been shown to inhibit the expression of the epidermal growth factor receptor mRNA (Witters *et al*, *Breast Cancer Res Treat* 53:41-50 (1999)) and 25-mer oligonucleotides have been shown to decrease the expression of adrenocorticotrophic hormone by greater than 90% (Frankel *et al*, *J Neurosurg* 91:261-7 (1999)). However, it is appreciated that it may be desirable to use oligonucleotides with lengths outside this range, for example 10, 11, 12, 13, or 14 bases, or 36, 37, 38, 39 or 40 bases.

Antisense polynucleotides may be administered systemically. Alternatively the inherent binding specificity of polynucleotides characteristic of base pairing is enhanced by limiting the availability of the polynucleotide to its intended locus *in vivo*, permitting lower dosages to be used and minimising systemic effects. Thus, polynucleotides may be applied locally to achieve the desired effect. The concentration of the polynucleotides at the desired locus is much higher than if the polynucleotides were administered systemically, and the therapeutic effect can be achieved using a significantly lower total amount.

The local high concentration of polynucleotides enhances penetration of the targeted cells and effectively blocks translation of the target nucleic acid sequences.

- 5 The antisense polynucleotides can be delivered to the locus by any means appropriate for localised administration of a drug. For example, a solution of the antisense polynucleotides can be injected directly to the site or can be delivered by infusion using an infusion pump. The antisense polynucleotides also can be incorporated into an implantable device which when placed  
10 adjacent to the desired site, to permit the antisense polynucleotides to be released into the surrounding locus.

The antisense polynucleotides may be administered via a hydrogel material. The hydrogel is non-inflammatory and biodegradable. Many such materials  
15 now are known, including those made from natural and synthetic polymers. In a preferred embodiment, the method exploits a hydrogel which is liquid below body temperature but gels to form a shape-retaining semisolid hydrogel at or near body temperature. Preferred hydrogel are polymers of ethylene oxide-propylene oxide repeating units. The properties of the polymer are  
20 dependent on the molecular weight of the polymer and the relative percentage of polyethylene oxide and polypropylene oxide in the polymer. Preferred hydrogels contain from about 10% to about 80% by weight ethylene oxide and from about 20% to about 90% by weight propylene oxide. A particularly preferred hydrogel contains about 70% polyethylene oxide and 30%  
25 polypropylene oxide. Hydrogels which can be used are available, for example, from BASF Corp., Parsippany, NJ, under the tradename Pluronic<sup>R</sup>.

In this embodiment, the hydrogel is cooled to a liquid state and the oligonucleotides are admixed into the liquid to a concentration of about 1

mg polynucleotides per gram of hydrogel. The resulting mixture then is applied onto the surface to be treated, for example by spraying or painting during surgery or using a catheter or endoscopic procedures. As the polymer warms, it solidifies to form a gel, and the antisense polynucleotides  
5 diffuse out of the gel into the surrounding cells over a period of time defined by the exact composition of the gel.

The polynucleotides can be administered by means of other implants that are commercially available or described in the scientific literature, including  
10 liposomes, microcapsules and implantable devices. For example, implants made of biodegradable materials such as polyanhydrides, polyorthoesters, polylactic acid and polyglycolic acid and copolymers thereof, collagen, and protein polymers, or non-biodegradable materials such as ethylenevinyl acetate (EVAc), polyvinyl acetate, ethylene vinyl alcohol, and derivatives  
15 thereof can be used to locally deliver the polynucleotides. The polynucleotides can be incorporated into the material as it is polymerised or solidified, using melt or solvent evaporation techniques, or mechanically mixed with the material. In one embodiment, the polynucleotides are mixed into or applied onto coatings for implantable devices such as dextran coated  
20 silica beads, stents, or catheters.

The dose of polynucleotides is dependent on the size of the polynucleotides and the purpose for which is it administered. In general, the range is calculated based on the surface area of tissue to be treated. The effective dose  
25 of polynucleotides is somewhat dependent on the length and chemical composition of the polynucleotides but is generally in the range of about 30 to 3000  $\mu\text{g}$  per square centimetre of tissue surface area.

The polynucleotides may be administered to the patient systemically for cosmetic, therapeutic and prophylactic purposes. The polynucleotides may be administered by any effective method, for example, parenterally (eg intravenously, subcutaneously, intramuscularly) or by oral, nasal or other means which permit the oligonucleotides to access and circulate in the patient's bloodstream. Polynucleotides administered systemically preferably are given in addition to locally administered polynucleotides, but also have utility in the absence of local administration. A dosage in the range of from about 0.1 to about 10 grams per administration to an adult human generally will be effective for this purpose.

It will be appreciated that antisense agents also include larger molecules which bind to said Kidins220 mRNA or genes and substantially prevent expression of said Kidins220 mRNA or genes and substantially prevent expression of said protein. Thus, expression of an antisense molecule which is substantially complementary to said Kidins220 mRNA is envisaged as part of the invention.

The said larger molecules may be expressed from any suitable genetic construct and delivered to the patient. Typically, the genetic construct which expresses the antisense molecule comprises at least a portion of the said Kidins220 cDNA or gene operatively linked to a promoter which can express the antisense molecule in the cell.

Although genetic constructs for delivery of polynucleotides can be DNA or RNA it is preferred if it is DNA. Equivalent genetic constructs can be used to deliver antisense polynucleotides to a patient as described above in relation to the delivery of polynucleotides encoding Kidins220 polypeptide or variants thereof.



Preferably, the genetic construct is adapted for delivery to a human cell.

Preferably, the antisense molecule is one which is capable of decreasing or  
5 inhibiting the function of the Kidins220 gene. Any decrease in the function  
of the said gene, for example the expression of the gene, is preferably a  
reduction of at least 10%, 20%, 30%, 40%, 50% of the level observed.  
More preferably, the function is reduced by 60%, 70%, 80% or 90%.  
Preferably, the reduction in function is such that the Kidins220 gene is  
10 expressed to the same level as observed in tissues in the same patient which  
are not cancerous or pre-cancerous.

In a preferred embodiment, the polynucleotide which is antisense further  
comprises a vector which is designed to express antisense DNA. Hence, the  
15 invention further provides a polynucleotide comprising a nucleic acid  
sequence which is antisense to a polynucleotide encoding the Kidins220  
polypeptide for use in medicine, especially in the manufacture of a  
medicament for treating cancer.

20 Further examples of agents which inhibit the function of the Kidins220 gene  
or its product include antibodies capable of selectively binding the  
Kidins220 polypeptide or a variant thereof, and ribozyme molecules which  
selectively cleave polynucleotides encoding the Kidins220 polypeptide.  
Suitable antibodies include those of the invention as described above which  
25 interfere with the function of the Kidins220 polypeptide, for example, by  
binding the serine at residue 919 (or its equivalent) and preventing  
phosphorylation or dephosphorylation, or by binding a protein interaction  
domain and thereby preventing a protein-protein interaction required for a  
function of the Kidins220 polypeptide. Methods of producing ribozymes

capable of selectively cleaving polynucleotides are well known in the art. Suitable ribozymes can be developed with only routine experimentation.

5 The cancer referred to in any one of the preceding methods of diagnosing, prognosing or treating cancer or monitoring the progression of cancer in a patient may be any type of neoplasm or tumour. Preferably, the cancer is one of astrocytic origin, such as glioblastoma multiforme, or the cancer is a prostate cancer. The level of Kidins220 expression is known to be especially elevated in cells from these types of cancers.

10

The invention further provides an agent capable of inhibiting the function of the Kidins220 gene or its product, an antibody according to the invention, and a polynucleotide which is antisense to a polynucleotide according to the first or third aspects of the invention, for use in medicine.

15

The invention also provides the use of an agent capable of inhibiting the function of the Kidins220 gene or its product in the manufacture of a medicament for treating cancer. Preferably, the agent is an antibody according to the invention or a polynucleotide which is antisense to a polynucleotide according to the first or third aspects of the invention.

20

The invention further provides the use of an agent which is capable of detecting the expression of the Kidins220 gene in the manufacture of a diagnostic reagent for diagnosing or prognosing cancer or for monitoring the progression of cancer in a patient. Preferably, the agent is an antibody according to the invention or a polynucleotide according to the first aspect of the invention.

25

The uses discussed above may be in respect of any cancer. Preferably, the cancer is glioblastoma multiforme or prostate cancer.

A further aspect of the present invention provides compounds comprising  
5 (i) a moiety which selectively binds the Kidins220 polypeptide and (ii) a further moiety.

According to a preferred embodiment, the moiety capable of selectively binding to the Kidins220 polypeptide is an antibody, preferably an antibody  
10 of the present invention as described above.

In an alternative embodiment, the moiety capable of selectively binding to the Kidins220 polypeptide is a peptide. As discussed above, it is known from Kong *et al* (2001) *J. Neurosci.* 21:176-185 that Kidins220 physically  
15 associates with the TrkA and p75 neurotrophin receptors.

Any cognate ligand for Kidins220 which is capable of selectively binding a region of the polypeptide which is located extracellularly may be useful. The extracellular regions of Kidins220 are likely to be located within the  
20 residues located between transmembrane domains 1 and 2 and between domains 3 and 4. These extracellular regions may be represented by amino acids 521-525 and 661-688 of the Kidins220 polypeptide sequence given in Figure 1b, or as described in Kong *et al* (2001) (*supra*). Any cognate ligands for Kidins220 will be a suitable moiety for selectively binding  
25 Kidins220. Peptides binding Kidins220 can be identified by means of a screen. A suitable method or screen for identifying peptides or other molecules which selectively bind Kidins220 may comprise contacting the Kidins220 polypeptide with a test peptide or other molecule under conditions where binding can occur, and then determining if the test

molecule or peptide has bound Kidins220. Methods of detecting binding between two moieties are well known in the art of biochemistry. Preferably, the known technique of phage display is used to identify peptides or other ligand molecules which bind to Kidins220. An alternative  
5 method includes the yeast two hybrid system.

Peptides or other agents which selectively bind the Kidins220 polypeptide include those which modulate or block the function of Kidins220.

10 Suitable peptides may be synthesised as described in more detail above.

The further moiety may be any further moiety which confers on the compound a useful property with respect to the treatment or imaging or diagnosis of diseases. Such diseases or other conditions or states are  
15 described in more detail below. In particular, the further moiety is one which is useful in killing tumour cells or imaging cancer. Preferably, the further moiety is one which is able to kill the cancer cells to which the compound is targeted.

20 In a preferred embodiment of the invention the further moiety is directly or indirectly cytotoxic. In particular the further moiety is preferably directly or indirectly toxic to cells in a tumour or cells which are in close proximity to and associated with a tumour.

25 Further embodiments and examples of further moieties which may be useful in the compound of the invention are described in more detail in WO 00/64480 (incorporated herein by reference).

Alternatively, the compound may be produced as a fusion compound by recombinant DNA techniques whereby a length of DNA comprises respective regions encoding the two moieties of the compound of the invention either adjacent one another or separated by a region encoding a linker peptide which does not destroy the desired properties of the compound. Conceivably, the two portions of the compound may overlap wholly or partly.

The DNA is then expressed in a suitable host to produce a polypeptide comprising the compound of the invention.

The invention also provides a kit of parts (or a therapeutic system) comprising (1) a compound of the invention wherein the further moiety which is able to convert a relatively non-toxic prodrug into a cytotoxic drug and (2) a relatively non-toxic prodrug. The kit of parts may comprise any of the compounds of the invention and appropriate prodrugs as herein disclosed.

The invention also provides a kit of parts (or a therapeutic system) comprising (1) a compound of the invention wherein the further moiety is able to bind selectively to a directly or indirectly cytotoxic moiety or to a readily detectable moiety and (2) any one of a directly or indirectly cytotoxic or a readily detectable moiety to which the further moiety of the compound is able to bind.

In a further embodiment, the binding moiety may be comprised in a delivery vehicle for delivering nucleic acid to the target. The delivery vehicle may be any suitable delivery vehicle.

Alternatively, the further portion may comprise a polypeptide or a polynucleotide encoding a polypeptide which is not either directly or indirectly cytotoxic but is of therapeutic benefit.

- 5 In a further embodiment of the invention, the further moiety comprised in the compound of the invention is a readily detectable moiety.

By a "readily detectable moiety" we include the meaning that the moiety is one which, when located at the target site following administration of the  
10 compound of the invention into a patient, may be detected, typically non-invasively from outside the body and the site of the target located. Thus, the compounds of this embodiment of the invention are useful in imaging and diagnosis.

- 15 The compounds of the final aspect of the invention are useful in treating, imaging or diagnosing disease, particularly cancer and neurodegenerative conditions.

Thus, the invention further provides a method of imaging cancer in the body  
20 of an individual the method comprising administering to the individual an effective amount of a compound according to the invention wherein the further moiety is a readily detectable moiety.

In a preferred embodiment of this aspect of the invention, the method of  
25 imaging cancer in an individual comprises the further step of detecting the location of the compound in the individual.

Detecting the compound or antibody can be achieved using methods well known in the art of clinical imaging and diagnostics. The specific method

required will depend on the type of detectable label attached to the compound or antibody. For example, radioactive atoms may be detected using autoradiography or in some cases by magnetic resonance imaging (MRI) as described above.

5

The invention further provides a method of diagnosing or prognosing cancer in an individual the method comprising administering to the individual an effective amount of a compound according to the invention wherein the further moiety is a readily detectable moiety.

10

The method may be one which is an aid to diagnosis.

In a preferred embodiment of this aspect of the invention, the method of diagnosing, or aiding diagnosis of, cancer in an individual comprises the further step of detecting the location of the compound in the individual.

15

The invention still further provides a method of treating an individual in need of treatment, the method comprising administering to the individual an effective amount of a compound according to the invention wherein the further moiety is a cytotoxic or therapeutic moiety.

20

Typically in the above-mentioned methods of treatment, the further moiety is one which destroys or slows or reverses the growth of the cancer.

For example, in the case where the compound of the invention comprises a readily detectable moiety or a directly cytotoxic moiety, it may be that only the compound, in a suitable formulation, is administered to the patient. Of course, other agents such as immunosuppressive agents and the like may be administered.

25

In respect of compounds which are detectably labelled, imaging takes place once the compound has localised at the target site.

- 5    However, if the compound is one which requires a further component in order to be useful for treatment, imaging or diagnosis, the compound of the invention may be administered and allowed to localise at the target site, and then the further component administered at a suitable time thereafter.
- 10   In a preferred embodiment of these aspects of the invention, the compounds are suitable for use in medicine.

15   The invention further contemplates a pharmaceutical composition comprising a compound of the invention and a pharmaceutically acceptable carrier.

20   A compound of the invention may be used in the manufacture of an agent for imaging cancer in a body of an individual or a diagnostic or prognostic agent for cancer or in the manufacture of a medicament for treating cancer.

The invention will now be described in more detail with reference to the following Figures and Examples:



**Figure 1.****Cloning and tissue distribution of Kidins220, a novel PKD-interacting protein.**

- (A) Immunoprecipitation with anti-PKD/PKC $\mu$  antibodies was carried out on PC12 cells treated (+) or untreated (-) with 200 nM PDB and the immunocomplexes were tested in an *in vitro* kinase assay as described in Experimental Procedures in Example 1. The samples were analysed by SDS-PAGE and autoradiography. The radioactive bands corresponding to PKD and to a 220 kDa protein (p220) are indicated.
- (B) Primary sequence of Kidins220. The predicted protein sequence corresponding to the complete ORF of Kidins220 cDNA is shown. The three peptides obtained by mass spectroscopy are double underlined and correspond to EST sequences AA998517 (rat), N83544 (human) and AA685317 (rat), respectively. The N-terminal 11 ankyrin repeats (AA 37-69, 70-102, 103-135, 137-169, 170-202, 203-235, 236-268, 269-301, 302-334, 335-367, 368-402; grey boxes), the ATP-binding motif (AA 467-474; bold underlined), the four transmembrane domains (AA 500-520, 526-546; 660-680 and 689-709; dark boxes), serine 919 (circled), and the splicing cassette (arrowheads) are indicated. The nucleotide sequence data are to be made available from GenBank/EMBL/DDBJ under the accession number AF239045.
- (C) Schematic representation of the structure of rat Kidins220 and similarity (grey boxes) with the *C. elegans* (F36H1.2) and the *D. melanogaster* (AC007175) orthologues.
- (D) Northern blots with 2  $\mu$ g of poly(A)<sup>+</sup> RNA from each of the indicated tissues (left) and 10  $\mu$ g of total RNA from each cell line (right) were hybridised with a random-primed full length Kidins220 cDNA probe. Sizes of RNA markers (in kb) are indicated.

**Figure 2.****Kidins220 is an ATP-binding integral membrane protein**

- (A) 50 µg of PC12 lysates were analysed by Western blot with an antibody against a carboxy-terminal Kidins220 peptide (AA 1747-1763). The antibody was used directly in Western blot, either without (-) or with (+) pre-incubation with the immunising carboxy-terminal peptide or in immunoprecipitation followed by Western blot (ip). The grey arrowhead indicates the position of the IgG heavy chain.
- (B) Kidins220 and PKD immunoprecipitates were incubated in the dark (-) or under u.v. light (+) with 8-azido [<sup>32</sup>P] α-ATP as described in the Experimental Procedures in Example 1, before SDS-PAGE analysis and autoradiography. Major radiolabelled bands corresponding to Kidins220 and PKD are indicated.
- (C) Analysis on sucrose velocity gradients of Kidins220 and PKD distribution in post-nuclear supernatants from PC12 cells. Fractions (1-18), from the top to the bottom, were probed in Western blot for endogenous Kidins220 and PKD. The two major peaks (low density, 4-8 and high density, 11-15) containing Kidins220 were further fractionated on two separated sucrose equilibrium gradients. Fractions were analysed by Western blot as above. Despite the fact that the majority of PKD is soluble in unstimulated cells, a sub-pool of PKD co-migrates with both Kidins220 peaks.
- (D) PC12 cells homogenate (SF), nuclear pellet (NP), post-nuclear supernatant (PNS) and the correspondent soluble (S) and membrane (P) fractions were prepared as described in the Experimental Procedures of Example 1. Equivalent amounts of membrane pellet (100 µg) were incubated on ice in a buffer containing either 1 M KCl (KCl), or 100 mM sodium carbonate (pH11). Samples were centrifuged and both the soluble

(S) and the pellet (P) fractions were analysed by Western blot with anti-Kidins220 antibodies.

(E) Triton-X 114 partitioning results in the distribution of Kidins220 signal in the aqueous (Aq) phase or the insoluble pellet (P), whilst only a minor  
5 fraction associates with the detergent (D) phase as shown by Western blot with anti-Kidins220 antibodies.

### Figure 3.

**Kidins220 is a plasma membrane protein present in lipid rafts.**

10 (A) Immunofluorescence analysis of PC12 cells was performed by using the Kidins220 carboxy-terminal polyclonal antibody in the absence (a) or presence (b) of the immunising peptide. Specific Kidins220 signal was observed both on the plasma membrane (arrowheads) and on an intracellular compartment. In non-permeabilising conditions (np, panel c),  
15 no specific staining was observed. Upon NGF-differentiation, Kidins220 immunostaining concentrates at the tips of the neurites (panel d, arrowheads and inset). Confocal microscopy images corresponding to the projection of a series of 0.4µm sections are shown. Bar: 5 µm. Immuno electron microscopy analysis of Kidins220 localisation in a representative section  
20 from a pellet of prefixed NGF-differentiated PC 12 cells is shown in panel e. Kidins220 is distributed predominantly on the plasma membrane of neurite extensions.

(B) PC12 membranes were incubated on ice for 30 min in a buffer containing either 1% Triton-X100 (TX100) or 1% saponin and 1% Triton  
25 X100 (saponin). After centrifugation, the solubilised proteins (S) and an insoluble pellet (P) were loaded and analysed by Western blot with anti-Kidins220 antibodies.

(C) NGF-differentiated PC12 cells were processed for immunofluorescence using anti-Kidins220 polyclonal antibodies (grey) and a monoclonal

antibody against the GPI-anchored protein Thy-1 (grey). Confocal microscopy images corresponding to a single section (0.4  $\mu\text{m}$ ) of a series are shown. The "Merge" panel shows the extent of colocalisation between the two proteins (white). This merged image corresponds to a single section  
5 taken at the level indicated by arrowheads in the top (z-x) and lateral (z-y) views. These cross-sections are derived from the complete series of Thy-1 and Kidins220 images. Scale: 5  $\mu\text{m}$ .

**Figure 4.**

10 **Kidins220 is a phosphoprotein interacting with PKD *in vitro* and *in vivo*.**

(A) PDB-treated (+) or untreated (-) PC12 cells were solubilized and total lysates were subjected to immunoprecipitation with PKD (PKD ip) or Kidins220 (Kidins220) polyclonal antibodies or the corresponding pre-immune serum (control ip). The antibodies alone (Ab) were also tested in  
15 the absence of cell lysates. Immunoprecipitates were separated by SDS-PAGE and analysed by Western blot with Kidins220 and PKD antibodies as indicated.

(B) Immunoprecipitates obtained as described in panel A were assayed in  
20 an *in vitro* kinase assay (IVK) (see Experimental Procedures in Example 1). Radioactive bands correspondent to Kidins220 and PKD are present in both anti-Kidins220 and anti-PKD immunoprecipitates. The asterisk labels the autophosphorylated PKD band.

25 **Figure 5.**

**PKD phosphorylates directly Kidins220 at a single Serine residue.**

(A) Vector alone (pcDNA3) or PKD mutants (PKD-Wt, PKD-D733A, PKD-S744E/S748E) and recombinant Kidins220-VSVG were transfected separately in COS-7 cells, stimulated with (+) or without (-) PDB and, after

lysis, treated with anti-PKD and anti-VSVG antibodies. PKD and Kidins220 beads were then mixed and incubated for 10 min at 30° C in an *in vitro* kinase assay. Samples were analysed by SDS-PAGE<sup>3</sup> and autoradiography. PKD-Wt and the constitutively active mutant  
5 S744E/S748E efficiently phosphorylate Kidins220, whilst the dominant negative mutant PKD-D733A is ineffective.

(B) Sequence of the peptides derived from Kidins220 sequence used in the *in vitro* exogenous substrate phosphorylation assays. Ser919 and its substitution by Ala919 are in bold. C-term corresponds to the carboxy-  
10 terminal region of Kidins220 (AA 1747-1763).

(C) PKD-Wt containing immuno-complexes derived from COS-7 transfected cells treated (+, filled bars) or untreated (-, empty bars) with 100 nM PDB were incubated with the different peptides (Syn-2, C-term, Ala 919 and Ser919) and the level of peptide phosphorylation was measured  
15 (Iglesias *et al.*, 1998c). Results are expressed as fold increase over Syn-2 phosphorylation by PKD under non-stimulated conditions and the values are the mean  $\pm$ S.E. of two independent experiments performed in triplicate.

(D) PKD-Wt and its mutants PKD-D733A and PKD-S744E/S748E were immunoprecipitated from COS-7 cells treated (+, filled bars) or untreated (-, empty bars) with 100 nM PDB and then incubated with 2.5 mg/ml of the Ser919 peptide in an *in vitro* kinase assay. The level of peptide phosphorylation was measured as specified in Experimental Procedures. The values shown are the mean  $\pm$ S.E. of two independent experiments performed in triplicate.  
20

(E) Effect of peptide substrate competition on PKD-mediated Kidins220 phosphorylation. PC12 cells treated (+) or untreated (-) with 100 nM PDB were lysed and endogenous Kidins220 and PKD were immunoprecipitated with anti-PKD antibodies. The immune-complexes were incubated in the presence of 2.5 mg/ml of the different peptides (C-terminal, Ala 919 and  
25

Ser919) in an *in vitro* kinase assay and then analysed by SDS-PAGE and autoradiography. Only the Ser919 peptide competes with Kidins220 phosphorylation.

5 **Figure 6.**

**Degradation of Kidins220 in cells expressing the PKD dominant negative mutant PKD-D733A.**

(A) PC12 cells transfected with GFP-PKD-Wt (panel a) or the dominant negative mutant GFP-PKD-D733A (panel b) were fixed and analysed by  
10 confocal microscopy. In contrast to the homogeneous distribution of GFP-PKD-Wt, the dominant negative mutant GFP-PKD-D733A accumulates in large aggregates (panel b and inset). Confocal microscopy images corresponding to the projection of a series of 0.4  $\mu$ m sections are shown. Panels a and b, bar 5  $\mu$ m; inset, bar 2.5  $\mu$ m.

15 (B) PC12 cells transfected with the dominant negative mutant GFP-PKD-D733A (panel a) were fixed and immunostained for endogenous Kidins220 (grey, panel b). The merged image is shown in panel c, where the white indicates colocalization. PKD and Kidins220 are both present in large aggregates (arrowheads). Confocal microscopy images correspond to the  
20 projection of a series of 0.4  $\mu$ m sections. Representative images of 100 cells are shown. Bar 5  $\mu$ m.

(C) Vector alone (-) or containing Kidins220-VSGV (Kidins220-VSVG) was co-transfected with pcDNA3 (-) or PKD-Wt (Wt) or its mutants PKD-D733A (D733A) or PKD-S744E/S748E (S744E/S748E) in COS-7 cells.  
25 After 72 hours, the level of Kidins220-VSVG and PKD mutants present in 100  $\mu$ g of total cell pellet was analysed by Western Blot with anti-VSVG antibodies and anti-PKD antibodies, respectively. Co-transfection with the dominant negative mutant PKD-D733 causes the disappearance of

Kidins220-VSVG, whilst the constitutively active PKD-S744E/S748E determines its accumulation.

**Figure 7.**

- 5    **The complete DNA sequence of the rat Kidins220 cDNA.** The open reading frame (ORF) corresponds to bases 1 to 5289, which encode amino acids 1 to 1763 as shown in Figure 1b.

**Figure 8.**

- 10   **Sequence of several splice variants of Kidins220 present in rat pheochromocytoma cells.**

Sequence analysis of independent Kidins220 clones highlighted the presence of a variable region in the central portion of the molecule. The  
15   identified cassette has a maximum dimension of 101 amino acid residues (clone 6) with inserts varying in dimension and sequence. Numbers without square brackets in this figure refer to the numbering shown in Figure 16 (amino acids) and Figure 7 (nucleotide sequence). Numbering shown within square brackets in this figure refer to the numbering shown in Figure  
20   18 (amino acids) and Figure 17 (nucleotide sequence).

**Figure 9.**

**Kidins220 is distributed in light and heavy plasma membrane fractions and co-localises with PKD sub-pools.**

- 25   (A) Analysis on sucrose velocity gradients of Kidins220 and PKD distribution in post-nuclear supernatants from PC12 cells. Fractions (1-18), from the top to the bottom, were probed in Western blotting for endogenous Kidins220 and PKD. The two major peaks (top fractions, 4-8 and lower fractions, 11-15) containing Kidins220 were further fractionated in (B) on

two separated sucrose equilibrium gradients (top, fractions 4-8; bottom, fractions 11-15). Fractions were collected (1-10) and analysed by Western blot as above. Despite the fact that the majority of PKD is soluble in unstimulated cells, a sub-pool of PKD co-migrates with both Kidins220 peaks.

**Figure 10.**

An alignment of the two human Kidins220 amino acid sequences which are publicly available from GenBank database entries AL133620 and AB033076. In the region of overlap where very low sequence identity is observed (residues 10006-1031 of AL133620 and residues 506-531 of AB033076), the sequence of AB033076 is considered to be correct by comparison with the sequence of rat Kidins220.

**Figure 11.**

The full length amino acid sequence of human Kidins220.

**Figure 12.**

Sequence comparison between different Kidins220 orthologues.

20

*Rattus norvegicus* (Rn), Human (Hs), *Drosophila* (Dm) and *C. elegans* (Ce) Kidins220 sequences were aligned with Clustal X using default parameters (gap open penalty: 10; gap extension penalty: 0.2; protein weight matrix: Gonnet). The position of invariable residues is indicated by asterisks, residues with similar charge, hydrophobic index or aromatic properties are indicated by a double dot, residues with the similar character (hydrophilic or hydrophobic) by a single dot.



**Figure 13.**

The cDNA sequence of human Kidins220 assembled from GenBank database entries AL133620 and AB033076. Nucleotides corresponding to the translated region are in capitals. Non-translated nucleotides are in lower case.

**Figure 14.**

A list of publicly available ESTs which display significant alignment to (a) rat and (b) human Kidins220. For both the rat and human sequences, the ankyrin repeat regions (shown in Figure 1b) were masked during analysis of sequence similarities, to exclude ESTs which are simply ankyrin repeat containing sequences.

**Figure 15.**

An alignment of the human (shown as the upper sequence) and rat (shown below the human sequence) Kidins220 cDNA sequences. The coding region is shown in capitals, and non-coding nucleotides are shown in lower case.

**Figure 16.**

The human Kidins220 promoter region located within Genbank Accession number AC012495 (human chromosome 2). The sequence between positions 34609-36632 denotes the region containing the putative transcription start site. Underlined positions 36510 to 36632 correspond to the 5' end of the human Kidins220 cDNA sequence that is shown in Figure 13. The sequence was derived by using the 5' end of the assembled human Kidins220 DNA sequence shown in Figure 13 to search the publicly available human genome sequence.

**Figure 17.**

**DNA sequence of rat Kidins220. The sequence corresponds to Genbank Accession No AF239045.**

**5 Figure 18.**

**Amino acid sequence of rat Kidins220. The sequence corresponds to GenBank Accession No AAG35185.**

**Figure 19.**

- 10 Representation of the domain structure of Kidins220, and the relative localisation of various motifs within the polypeptide:**

aa17 – 391 Ankyrin repeat region circular profile.

aa707-748 Domain of unknown function DUF38.

aa1246-1256 Na<sup>+</sup>/K<sup>+</sup> ATPase C-terminus.

- 15 aa1204-1224 PIPLC X DOMAIN:**

Potential phosphorylated amino acids are indicated by a encircled "P", DUF38 domains are indicated by an encircled "DUF38".

**Figure 20.**

- 20 Distribution of Kidins220 in developing mice embryos.**

**Figure 21.**

**Distribution of Kidins220 in the brain.**

- 25 Figure 22.**

**SAGE analysis of human Kidins220 expression.**

**Figure 23.**

**Immunohistochemical analysis of Kidins220 expression in the U87MG cell line.**

5 **Figure 24.**

**Nucleotide sequence comparison of the rat DNA sequence shown in Figure 7 with that shown in Figure 17. The sequence from Figure 7 is shown above the sequence from Figure 17.**

10 **Figure 25.**

**Amino acid sequence comparison of the rat amino acid sequence shown in Figure 1b with that shown in Figure 18. The sequence from Figure 1b is shown above the sequence from Figure 18.**

15 **Example 1: Kidins220, a new neuronal PKD substrate, defines a novel role for PKD in the regulation of protein turnover.****Purification and cloning of Kidins220, a novel PKD substrate.**

PKD is expressed in many tissues, although the brain has one of the highest  
20 mRNA levels (Valverde *et al.*, 1994). In the search for PKD-specific  
substrates, we selected the pheochromocytoma cell line PC12 because it has  
been extensively used to investigate multiple signal transduction pathways  
(Marshall, 1995). Additionally, PC12 cells, (rat cancer cells of  
neuroendocrine origin) represent a well established cellular model to study  
25 neurite outgrowth and regulated secretion and express considerable amounts  
of PKD (Wooten *et al.*, 1997).

PC12 lysates were immunoprecipitated with specific PKD antibodies  
followed by an *in vitro* kinase assay. Co-immunoprecipitating proteins that

appeared to be phosphorylated under stimulating conditions, such as phorbol ester treatment, were analysed by SDS-PAGE and autoradiography. A band of ~220kDa (p220), matching these requirements, was detected (Figure 1A). To isolate p220, extracts prepared from  $5 \times 10^{10}$  PC12 cells were subjected to the same steps of immunoprecipitation and *in vitro* kinase assay. After SDS-PAGE analysis, the resulting 220 kDa band was excised and directly processed for microsequencing. Protein fingerprinting and peptide sequence exhibited no significant sequence homology to any known protein. However, BLAST (National Center for Biotechnology Information, NCBI, Bethesda, MD) searches yielded a variety of human and rat expressed sequence tags (ESTs; peptide 1, FLFTDYNR, rat locus AA998517; peptide 2, TPSLSSLNSQDSSIEISK, human locus N83544; peptide 3, ANQNFDEIEGIR, rat locus AA685317; double underlined in Figure 1B).

The majority of p220 cDNA was amplified by RT-PCR and the remaining 5'-end was cloned by using 5'-RACE. The full-length clone was 7,140 bp and comprised 111 bp upstream of the initiation methionine. This 5' untranslated region contained several stop codons in all three reading frames and presented the correct sequence context for translational initiation (Kozak, 1992). The p220 clone had an extended 3'-untranslated region of 1,740 bp after the TGA stop codon. The open reading frame of 5,289 bp (Figure 7) encoded a putative protein of 1,763 amino acids (AA) with a predicted molecular mass of 196,561 kDa (Figure 1B). Because of its PKD interaction, we have named this protein Kidins220 (Kinase d interacting substrate of 220 kDa). Although human, rat and mouse ESTs corresponding to partial sequences of the Kidins220 cDNA were represented in the GenBank/EMBL/DDBJ database, analysis of protein databases using the BLAST Enhanced Alignment Utility algorithm failed to identify any

significant similarities to known proteins. However, BLAST searches found one positive match showing a high degree of homology (23.6% identity; 34.3% similarity) with a open reading frame present in the *C.elegans* cosmid F36H1.2 (GenBank accession number Q20109). BLAST searches  
5 revealed also several hits to an ordered segment of *D. melanogaster* chromosome 2R. The best hit was within a contiguous 188 Kb DNA region (GenBank accession numbers AC007175) from which it is possible to predict a protein with the same domain composition as observed in rat and *C. elegans* orthologues (Figure 1C). Sequence comparison of the rat and *C.*  
10 *elegans* clones revealed a higher degree of conservation in the amino-terminal region (residues 1-1065 presents 30% identity; 43.3% similarity), diverging more at the carboxy-terminal end (Figure 1C). During the preparation of this manuscript, two sequences corresponding to different fragments of the human Kidins220 homologue have been deposited to the  
15 GenBank (AB033076 and AL133620; Nagase *et al.*, 1999).

**Kidins220 is a multi-domain protein expressed in brain and neuronal cells.**

20 The primary amino acid sequence of Kidins220 can be subdivided into several structural and functional domains (Figure 1C). The amino-terminus begins with a region of ~400 residues comprising 11 ankyrin repeats, which are likely to be involved in protein-protein interactions (Sedgwick and Smerdon, 1999). Downstream of this motif, the sequence contains a  
25 consensus ATP/GTP binding site (P-loop) at residues 467-474 (AQWGSGKS), followed by four putative transmembrane domains. The presence of several membrane spanning domains strongly suggests that Kidins220 is an integral membrane protein, possibly localised to the plasma membrane (39.1% probability by using PSORT topology prediction

system). The similarity with the *C. elegans* and *D. melanogaster* orthologue is interrupted by a large insertion, which contains a proline-rich region (AA 984 to 1054), followed by a splicing cassette (AA 1140-1232). This last feature suggests that mammalian Kidins220 is the prototype of a family of membrane proteins constituted by several variants (Figure 8). The splicing cassette is followed by a short region in which the homology between the *C. elegans* and the rat protein returns (35.9% identity; 53.8% similarity). The carboxy-terminal half of the sequence is unrevealing, except for the abundance of serine and threonine residues, which could constitute potential phosphorylation sites for different kinases, including PKCs. Interestingly, a unique residue, Ser919, is present in a sequence context favourable for the specific phosphorylation by PKD (Figure 1B and Table I)(Nishikawa *et al.*, 1997).

TABLE I

## Consensus sequence for PKD phosphorylation

PEPTIDE	Position										
	-5	-4	-3	-2	-1	0	+1	+2	+3	+4	+5
$\mu$ -peptide	L	V	R	Q	M	S	V	A	F	F	F
KiDin220-S919	I	T	R	Q	M	S	F	D	L	T	K
Syntide 2	L	A	R	T	L	S	V	A	G	L	P

We examined the expression of Kidins220 mRNA in various rat tissues and different cell lines. Northern blot analysis with Kidins220 cDNA revealed the presence of a ~7.6 kb mRNA in brain (Figure 1D). A much lower hybridisation signal was present in heart and it was undetectable in other tissues (Figure 1D). Interestingly, an additional second messenger of ~5 kb was also present in brain and heart. Kidins220 mRNA was expressed predominantly in PC12 cells and to a lesser extent in the neuroblastoma line N2A and the glioma line C6 (Figure 1D) but was not detectable in the other

cell lines tested. Altogether, these results suggest that Kidins220 is primarily expressed in tissues of nervous origin.

### **Kidins220 is an ATP-binding plasma membrane protein**

5

To investigate the properties of Kidins220, polyclonal antibodies were raised against the last seventeen residues of its carboxy-terminus (AA 1747-1763). This antiserum recognised a 220 kDa protein in PC12 lysates (Figure 2A). Pre-absorption with the immunising peptide completely abolished the  
10 recognition of the 220 kDa band, demonstrating its specificity (Figure 2A). The ability of this antibody to immunoprecipitate Kidins220 was also tested in PC12 lysates. As shown in Figure 2A, a 220 kDa band was specifically recovered, as revealed by subsequent Western blot with the same Kidins220 antiserum.

15

The presence of a P-loop within the primary sequence of Kidins220 prompted us to examine if Kidins220 is indeed an ATP-binding protein. The ATP binding in Kidins220 immunoprecipitates was tested with 8-azido- $[^{32}\text{P}]\alpha\text{-ATP}$  (see Experimental procedures). Under these conditions,  
20 Kidins220 is covalently modified by the ATP analogue, as demonstrated by the radioactive band migrating at 220 kDa (Figure 2B). This labelling is specific, because it is absent in conditions not allowing the activation of the photoactivatable group (Figure 2B, no light) or in the presence of an excess of cold ATP. As expected, immunoprecipitation experiments with anti-  
25 PKD antibodies revealed two major radiolabelled bands corresponding to Kidins220 and PKD (Figure 2B, arrows). Phorbol ester treatment did not change the extent of labelling (data not shown), suggesting that Kidins220 phosphorylation by PKD is not regulating its ATP binding ability.

To study the predicted association of Kidins220 with cellular membranes, the postnuclear supernatant from undifferentiated PC12 cells was loaded on a sucrose velocity gradient (Dittie *et al.*, 1997) and fractions were analysed for specific markers. Kidins220 was absent from the first two fractions at the top of the gradient (Figure 2C), which are enriched mainly in cellular soluble proteins (Dittie *et al.*, 1997), suggesting that Kidins220 is membrane-bound. Kidins220 was distributed along the gradient mainly in two peaks (low density, fractions 4-7 and high density, fractions 11-15; Figure 2C). Western blot with an antibody against  $\text{Na}^+/\text{K}^+$ -ATPase, a typical marker of the plasma membrane, showed that Kidins220 and  $\text{Na}^+/\text{K}^+$ -ATPase overlap in the second peak representing the heavier fractions of the gradient (not shown). The presence of a lighter peak of Kidins220 suggests that it is also associated with a low-density lipid-enriched membrane population. We examined the distribution of PKD in the same gradient with anti-PKD antibodies. The vast majority of PKD signal (>95%) appeared in the fractions at the top of the gradient, confirming its mainly cytosolic localisation (Figure 2C). However, a low but significant PKD signal was also found in the remaining fractions (Figure 2C). The two Kidins220-enriched peaks were further fractionated in sucrose equilibrium gradients and protein distribution was analysed by Western blot. As shown in Figure 2C, PKD and Kidins220 signals overlapped in both low and high density peaks (fractions 4-5 and 6-8, respectively), suggesting that two PKD subpools colocalise with Kidins220 under basal conditions.

To determine whether Kidins220 is an integral membrane protein, PC12 cells were fractionated and the post-nuclear supernatant (PNS) was further centrifuged to give a soluble fraction (S) and a membrane pellet (P). The 220 kDa band was only present in the membrane pellet (Figure 2D). To investigate in more detail the association of Kidins220 with membranes, the



membrane pellet was incubated with salt or exposed to mild alkaline conditions, two treatments that remove extrinsic membrane proteins. Kidins220 was found associated to the pellet fraction under all conditions (Figure 2D), indicating that Kidins220 is an integral membrane protein. To confirm this observation, phase partition of the membrane pellet with Triton X-114 was performed and equal aliquots of the aqueous phase (Aq), detergent phase (D) and insoluble pellet (P) were analysed. Strikingly, a high percentage of Kidins220 partitioned in the aqueous phase, although some of the protein remained in the detergent phase, as well as in the insoluble pellet (Figure 2E). This is not unprecedented. Giantin, a large (~400 kDa) integral component of Golgi membranes (Linstedt and Hauri, 1993), has been observed to partition in the aqueous phase upon Triton X-114 extraction, presumably due to its extensive hydrophilic surface. The large hydrophilic carboxy-terminal half of Kidins220 could be responsible for this effect.

#### **Kidins220 concentrates at neurite tips and lipid rafts**

The cellular localisation of Kidins220 was examined by immunofluorescence in PC12 cells using anti-Kidins220 antibodies. As shown in Figure 3A, Kidins220 distribution revealed an intense discontinuous staining at the plasma membrane, brighter in some membrane domains, being somehow dispersed through the cell body with a considerably high asymmetrical perinuclear labelling (Figure 3A, panel a). This intracellular and perinuclear staining did not colocalise with the Golgi markers GM130 and giantin, with the endoplasmic reticulum markers p62/calnexin, or with the intermediate compartment marker Ergic53 (not shown). The Kidins220 signal was blocked either by pre-absorption with the generating peptide or by omitting the permeabilisation step (Figure 3A,

panels b and c, respectively), confirming, on one hand, the specificity of the antibody and, on the other, the intracellular localisation of the carboxy-terminus of Kidins220. This fact and the presence of four transmembrane domains support a topology for Kidins220 in which both the carboxy- and the amino-terminus are intracellular. In nerve growth factor (NGF) differentiated PC12 cells, Kidins220 accumulated at neurite tips, suggesting the possibility of its involvement in neurite extension or growth cone organisation and dynamics (Figure 3A, panel d and inset). Immuno electron microscopy analysis of Kidins220 distribution in NGF-treated cells confirmed its dual localisation on the plasma membrane of neurites and on an unidentified internal membrane-associated compartment (Figure 3A, panel e).

To further analyse KiDin220 cellular distribution, the postnuclear supernatant from undifferentiated PC12 cells was loaded on a sucrose velocity gradient (Dittie *et al*, 1997). After centrifugation, gradient fractions were collected and analysed for the presence of specific markers, Kidins220 was distributed along the entire gradient mainly in two peaks running from fraction 4-7 and 11-15 (Figure 9A). Western blotting with an antibody against Na<sup>+</sup>/K<sup>+</sup>-ATPase, a typical marker of the plasma membrane, showed that Kidins220 and Na<sup>+</sup>/K<sup>+</sup>-ATPase overlap in the second peak representing the heavier fractions of the gradient (11-15; not shown). The lighter peak of Kidins220 running at the top of the gradient suggests that this protein is also specifically present in a low-density membrane population, possibly constituted by a membrane fraction enriched in lipids. Since Kidins220 co-immunoprecipitated with PKD in vivo, we examined the distribution of PKD in the same gradient with anti-PKD antibodies. The vast majority of PKD signal (>99%) appeared in the first two fractions at the top of the velocity gradient (Figure 9A), which are

enriched mainly in cellular soluble proteins (Dittie *et al.*, 1997). However, a low but significant PKD signal was also found along all the remaining fraction of the gradient (Figure 9A). These results provide evidence that only a small portion of the total amount of PKD co-localised with  
5 Kidins220 under these experimental conditions. This conclusion is supported by the results obtained by further fractionating the two Kidins220-enriched peaks in sucrose equilibrium gradients, followed by Western blotting with anti-Kidins220 and anti-PKD antibodies. As shown in Figure 9B, the two proteins overlap only partially, showing colocalisation  
10 in fractions 4-5 (peak 1) and 6-8 (peak 2).

Some plasma membrane components, including various glycosylphosphatidylinositol (GPI)-anchored proteins, transmembrane proteins and signalling molecules, co-isolate in non-ionic detergent-resistant  
15 fractions of the plasma membrane, known as detergent-insoluble glycolipid-enriched domain (DIGs) or lipid rafts (Harder and Simons, 1997; Brown and London, 1998). To gain insights into the association of Kidins220 with specific plasma membrane domains, its solubility in non-ionic detergent was analysed. Remarkably, Kidins220 was resistant to extraction by Triton  
20 X-100 at 4°C (Figure 3D). This result, together with the presence of Kidins220 in low density fractions (Figure 2C), strongly suggests its localisation to DIGs. A number of studies have used cholesterol-binding agents such as saponin to facilitate DIGs solubility in non-ionic detergents (Ilangumaran and Hoessli, 1998). As expected for a protein associated with  
25 lipid rafts, Kidins220 was totally solubilised when saponin and Triton-X100 were combined (Figure 3B). Furthermore, double labelling of NGF-differentiated PC12 cells with antibodies against Thy-1, a GPI-anchored protein targeted to DIGs (Madore *et al.*, 1999), and Kidins220 demonstrated

that indeed these two proteins extensively colocalise in discrete regions of the plasma membrane (Figure 3C).

### ***In vivo* interaction between Kidins220 and PKD.**

5

To extend the analysis of PKD and Kidins220 interaction, we assayed the presence of PKD in Kidins220 immunoprecipitates. As shown in Figure 4A, immunoprecipitation of Kidins220 from PC12 lysates resulted in the co-immunoprecipitation of PKD. The use of the antibody alone (Ab) or the  
10 pre-immune serum did not result in any recovery of PKD in the sample, confirming the specificity of the interaction. Reciprocally, as expected from the first *in vitro* kinase experiment (Figure 1A), anti-PKD antibody co-immunoprecipitated Kidins220 and their association was not significantly altered upon PDB treatment (Figure 4A). The co-immunoprecipitated PKD  
15 was fully functional, as tested by *in vitro* kinase assay carried out on the anti-Kidins220 beads (Figure 4B, asterisk). These results strongly suggest that Kidins220 is associated constitutively with a sub-population of PKD molecules and PKD/Kidins220 complexes are independent of phorbol esters stimulation.

20

### **PKD directly phosphorylates Kidins220**

Preliminary immunoprecipitation and *in vitro* kinase assays suggested that Kidins220 might be a potential substrate for PKD. In these assays, the 220  
25 kDa band was phosphorylated and the level of modification appeared to be increased by phorbol esters treatment (Figure 1A). Due to its richness in serine and threonine residues (9.4% and 6.5%, respectively) and the presence of multiple potential phosphorylation sites in the carboxy-terminal portion, we tested whether Kidins220 is a phosphoprotein *in vivo*.

Immunoprecipitation of extracts from [ $^{32}\text{P}$ ]-labelled PC12 cells with anti-Kidins220 and anti-PKD antibodies revealed a radioactive band of 220 kDa corresponding to Kidins220, which strongly suggests that Kidins220 is phosphorylated *in vivo* (not shown).

5

*In vitro* phosphorylation experiments were carried out by using recombinant Kidins220 as substrate. For this purpose, expression vectors containing a VSVG-tagged version of Kidins220 (Kidins220-VSVG) and PKD (Wt) were transfected separately in COS-7 cells and then immunopurified with  
10 anti-VSVG and anti-PKD antibodies. In the experiment presented in Figure 5A, immobilised Kidins220-VSVG was incubated with PKD-beads in the presence of [ $^{32}\text{P}$ ]  $\alpha$ -ATP, followed by SDS-PAGE and autoradiography. Under these conditions, a radioactive band corresponding to Kidins220 was detected and its level of phosphorylation increased following *in vivo*  
15 stimulation of PKD with phorbol ester (Figure 5A). Therefore, Kidins220 is specifically recognised and phosphorylated by PKD. To demonstrate conclusively that PKD kinase activity was directly responsible for Kidins220 phosphorylation, the constitutively active double mutant, PKD-S744E/S748E (S744E/S748E), and the dominant negative mutant, PKD-D733A (D733A), were used in parallel. As shown in Figure 5A, Kidins220  
20 was phosphorylated only when the active mutant was used and this phosphorylation completely disappeared in the presence of the dominant-negative PKD mutant. Altogether, these findings demonstrate that Kidins220 is phosphorylated *in vitro* by PKD, and strongly suggest that  
25 Kidins220 is most likely a PKD substrate *in vivo*.

To date, specific physiological substrates for PKD have not been identified. The only information available is derived from the use of synthetic peptides, such as syntide-2 (Valverde *et al.*, 1994; Van Lint *et al.*, 1995), which are

efficiently phosphorylated *in vitro*. More recently, a peptide library study predicted the optimal substrate sequence for PKD/PKC $\mu$  ( $\mu$ -peptide) and showed the  $\mu$ -peptide to be extremely specific for this isoform (Nishikawa *et al.*, 1997). Sequence analysis of Kidins220 revealed a consensus motif for PKD phosphorylation at position 919 (Ser919; Table I). On this basis, a synthetic peptide corresponding to this sequence was used as substrate for PKD (Kidins220-Ser919, AA 908-924) and the efficiency of phosphorylation compared to that of syntide-2. A mutated Ser to Ala peptide (Kidins220-Ala919) and the Kidins220 carboxy-terminal peptide (Kidins220 C-Term, AA 1747-1763) served as negative controls (Figure 5B). The Kidins220-Ser919 peptide was phosphorylated very efficiently by PKD-Wt and its phosphorylation was increased following *in vivo* activation of PKD by phorbol esters. The single substitution of Ser to Ala or the use of the carboxy-terminal peptide abolished the phosphorylation completely (Figure 5C). The level of Kidins220-Ser919 phosphorylation by unstimulated PKD was consistently twofold higher when compared to that of syntide-2 (Figure 5C), demonstrating that this peptide is a better PKD substrate and suggesting that this sequence may be specifically phosphorylated by PKD *in vivo*. To further support a direct phosphorylation of Kidins220 by PKD and to exclude the involvement in this process of other co-immunoprecipitating kinases, the PKD dominant negative mutant, PKD-D733A and a constitutively active mutant, PKD-S744E/S748E were tested in parallel. As shown in Figure 5D, Kidins220-Ser919 was fully phosphorylated in the presence of the constitutively active PKD mutant independently of phorbol ester stimulation. In contrast, no phosphorylation was observed when the dominant negative form of PKD was used. In addition, after co-immunoprecipitation of native PKD and Kidins220 from PC12 cells, the KiDin-Ser919 peptide was able to compete almost completely (>80%) the phosphorylation of endogenous Kidins220, whereas

the alanine mutant version or the carboxy-terminal peptide did not inhibit this process (Figure 5E).

### **PKD is a modulator of Kidins220 trafficking and turnover**

5

To provide evidence of the involvement of PKD in Kidins220 cellular function(s), we transfected different GFP-PKD mutants into PC12 cells and analysed their effects on endogenous Kidins220 by confocal microscopy. GFP-PKD-Wt in NGF-treated PC12 cells was uniformly distributed  
10 throughout the cell soma and neurites (Figure 6A, panel a) and was indistinguishable from the constitutively active mutant GFP-PKD-S744E/S748E (not shown). Strikingly however, the dominant negative GFP-PKD-D733A was mainly restricted to large intracellular aggregates (Figure 6A, panel b and inset), suggesting that PKD activity is essential for  
15 its physiological localisation in PC12 cells. Further analysis of GFP-PKD-D733A expressing cells revealed a partial colocalisation with Kidins220 in these large aggregates, as shown by the yellow pseudocolour in the merged confocal image (Figure 6B, arrowheads). In contrast, Kidins220 immunostaining in GFP-PKD-Wt or GFP-PKD-S744E/S748E cells was  
20 identical to the pattern observed prior to transfection (not shown), indicating that PKD over-expression itself does not alter Kidins220 distribution.

Unexpectedly, the level of PKD dominant negative expression seemed to inversely correlate with Kidins220 signal (Figure 6B, upper cell vs. lower  
25 cell in panel a and b). In order to gather biochemical evidence of the possible participation of PKD activity on Kidins220 protein levels, PKD-Wt, PKD-D733A and PKD-S744E/S748E were co-transfected with Kidins220-VSVG into COS-7 cells and total cell pellets were analysed by SDS-PAGE and immunoblotted with anti-PKD and anti-VSVG antibodies.

As shown in Figure 6C, Kidins220-VSVG was present in all samples but the one over-expressing the dominant negative mutant PKD-D733A. Moreover, when the constitutively active mutant PKD-S744E/S748E was present, the signal of Kidins220-VSVG was increased (Figure 6C). The  
5 levels of expression of the three different PKD forms (PKD-Wt, PKD-D733A and PKD-S744E/S748E) were similar, as assessed by Western blot with specific anti-PKD antibodies (Figure 6C, PKD). These results strongly indicate that the degradation of the Kidins220-VSVG is not a consequence of over-expression and/or tagging of the recombinant protein, but reflects a  
10 functional effect of PKD activity on Kidins220 stability. Inhibition of PKD activity, by means of the dominant negative mutant PKD-D733A, leads to the specific aggregation and degradation of Kidins220, providing evidence that PKD phosphorylation of its protein substrate(s) might be a key component in the modulation of their intracellular traffic and turnover.

15

## EXPERIMENTAL PROCEDURES

### Cell culture and transfection

Different cell lines: PC12, N2a, C6, U251, COS-7, NIH-3T3, HeLa and CHO, were obtained from the ICRF Cell Collection. When required, PC12  
20 cells were treated with NGF (50 ng/ml) for 5 days. PC12 cells were plated at 50-60% confluence on glass coverslips and, after 24 hours, transfection of 3 µg of plasmid DNA was performed using the TransFast™ reagent (Promega, WI). 48 hours after transfection and NGF treatment (50 ng/ml), cells were fixed and processed for immunofluorescence. COS-7 cells were  
25 transfected using Lipofectin® reagent (Life Technologies, UK) according to manufacturer specifications and were used 72 h later.



### Purification of Kidins220

PC12 cells were lysed in RIPA-buffer (25 mM Tris-HCl, pH 7.6, 1% Triton X-100, 1% sodium deoxycholate, 0.1% SDS, 150 mM NaCl, 2 mM EDTA, 2 mM dithiothreitol) with protease inhibitors (1 mM 4-(2-aminoethyl)-benzenesulfonyl fluoride and 10 µg/ml of aprotinin, leupeptin and pepstatin) for 60 min at 4°C. Extracts were immunoprecipitated with an anti-PKD-PKCµ antibody (D-20, sc-935, Santa Cruz Biotechnology, CA) for 3 hours at 4°C. The immunoprecipitate was loaded in a preparative 8% SDS-PAGE gel and stained with the Zinc-method (Bio-Rad, CA). The protein band was cut, trypsin digested for 8 hours at 37° C and extracted for 20 minutes with 50% acetonitrile in 5% trifluoroacetic acid. Peptides were derivatised with N-succinimidyl-2-morpholine acetate and *de novo* sequenced by low-energy collision-activated dissociation using a LCQ ion-trap MS (ThermoQuest, CA) fitted with a nanoelectrospray source, as previously described (Hoss *et al.*, 1999).

### Cloning of rat Kidins220 cDNA.

To clone most of the Kidins220 cDNA sequence, we used reverse transcriptase PCR (RT-PCR) using as template mRNA obtained from PC12 cells. Forward and reverse primers were designed based on EST sequences and overlapping RT-PCR clones were fused to generate a partial sequence, which covered 70% of the final Kidins220 clone. To obtain the full-length cDNA, 5'-RACE reactions were performed with the Marathon™ cDNA amplification kit (Clontech, UK). The absence of genomic DNA in the RNA samples was confirmed by the lack of PCR products when no reverse transcriptase was used. PCR products were subcloned in the pGEM-T Easy vector (Promega, WI) and at least three clones from each RT-PCR or each 5'-RACE reactions were sequenced on both strands. The complete DNA sequence of the rat Kidins220 clone has been deposited at GenBank with

the accession number AF239045. Several additional Kidins220 clones representing alternative splice variants were isolated and will be described in Example 2. Analysis of Kidins220 sequence has been performed using the following applications (intron-exon boundaries assignment by  
5 GENSCAN, <http://CCR-081.mit.edu/GENSCAN.html>; topology prediction by PSORT, <http://www.psort.nibb.ac.jp>; domain composition by Pfam 4.4, [http://www.sanger.ac.uk/Software/Pfam](http://www.sanger.ac.uk/Software/Pfam;);) )

### RNA preparation and Northern Blot analysis

10 Total RNA was prepared using TRIzol Reagent (Life Technologies, UK) following the manufacturer instructions. Northern blots were performed on nylon membranes (Nytran; Schleicher and Schuell, D) according to standard protocols. The full length Kidins220 cDNA was labelled and used to screen Northern blots containing 2 µg of poly(A)+ RNA purified from different rat  
15 tissues (Clontech, CA) or 10µg of total RNA from different cell lines. Equal loading between different lanes was assessed by methylene blue staining of 18S and 28S rRNAs.

### Eukaryotic expression

20 The coding sequence of rat Kidins220 was cloned in a modified version of pcDNA3.1/V5-HisA (Invitrogen, CA) obtained by insertion of the oligonucleotide 5'-TCG ACA CCC GGG AAG TTA ACG AGT CTA GAT CCA TCG ATC ACC GGT GT-3' at the *Xho* I/Age I site. The amino-terminal *Sma* I/*Xba* I fragment of Kidins220 was cut from pGEM-T-Easy  
25 and subcloned in the new polylinker. The carboxy-terminal half of Kidins220, starting from the unique *Xba*I site, was reamplified using a reverse oligonucleotide with a point mutation in the stop codon and a *Cla*I site (5'-ATC GAT CTT TTC ACT TCT CCT TCC TGA GAG A-3'). This PCR product was digested with *Xba*I/*Cla*I and sub-cloned in the new vector

to obtain Kidins220 complete cDNA followed by a His<sub>6</sub>-tag motif (pcDNA3-Kidins220-HisA). We added a Vesicular Stomatitis Virus Glycoprotein (VSVG) epitope sequence by introducing in the *Clal/AgeI* the oligonucleotide 5'-CGA TTA CAC CGA TAT AGA GAT GAA CAG GCT  
5 GGG AAA GTC A-3' (pcDNA3-Kidins220-VSVG-HisA). cDNA constructs containing wild-type and mutant PKD sequences and their GFP tagging have been previously described (Iglesias *et al.*, 1998c; Matthews *et al.*, 1999).

#### 10 Immunoprecipitation, *in vitro* Kinase and ATP-binding assays

Immunoprecipitation was performed with the desired primary antibodies and protein-A Sepharose beads for 3 hours at 4°C. Immune complexes were washed three times and analysed by SDS-PAGE and Western blot. Autophosphorylation and peptides phosphorylation were determined by in  
15 vitro kinase (IVK) assays as previously described (Iglesias *et al.*, 1998c). Syntide-2, Kidins220-Ser919 (RQMQRITRQMSFDLTK, AA 908-924), Kidins220-Ala919 (RQMQRITRQMAFDLTK) and Kidins220-C-terminal (CAASSESTGFGEERSIL, AA 1747-1763) peptides were used at a final concentration of 2.5 mg/ml. For ATP-binding assays, 8-azido [<sup>32</sup>P] α-ATP  
20 (2μCi/sample) was added to the immunoprecipitates in the dark and subsequently photoactivated for 20 min on ice with a long-wave UV lamp (Vilber Lourmat, F) prior SDS-PAGE and autoradiography.

#### Subcellular fractionation

25 PC12 cellular fractionation was performed as previously described (Dittie *et al.*, 1997). Membrane pellets were recovered from post-nuclear supernatants by spinning at 100,000xg for 30 min. To study Kidins220 membrane properties, membrane fractions were resuspended in breaking buffer (250 mM sucrose, 10 mM Hepes, 1 mM EDTA, 1 mM MgOAc plus protease

inhibitors) containing either 1 M KCl or 100 mM sodium carbonate, and left at 4°C for 30 min prior ultracentrifugation. Triton X-114 partitioning was done as described (Bordier, 1981). Kidins220 solubility in 1% Triton X-100, with or without 1% saponin, was tested in membrane pellets by incubation for 30 min at 4°C. Samples were centrifuged and fractions analysed in SDS-PAGE gel and Western blot. For peptide competition, the Kidins220 antiserum was pre-incubated with 5 mg/ml of immunising peptide for 30 min.

#### 10 Immunocytochemistry and electron microscopy

Cells grown on coverslips were either directly fixed and permeabilized in methanol at -20°C for 4 min, or, in the case of Thy-1, fixed for 4 min in 4% paraformaldehyde in PBS at 8°C, followed by methanol for 5 min at -20°C. Samples were incubated with the corresponding primary antibodies for 1 h at room temperature, detected with the suitable fluorophore-conjugated IgGs before mounting in Mowiol 4-88 (Harco, UK). The antibodies used in these experiments include: rabbit polyclonal antibody against the Kidins220 CAASSESTGFGEERESIL (AA 1747-1763) peptide, affinity-purified rabbit polyclonal antibody to rat GM130; rabbit polyclonal antibody against p62; monoclonal anti-Thy-1 (clone OX-7; Bukovsky *et al.*, 1983); monoclonal anti-Ergic53, monoclonal anti-VSVG epitope (P5D4; Soldati and Perriard, 1991), Texas Red- and Alexa A488-conjugated secondary antibodies were purchased from Molecular Probes (Eugene, OR). Raw images acquired by confocal microscopy (Zeiss LSM510, D) with a x63 Plan-Apochromatic oil objective were normalised for each colour separately. Confocal images presented are two dimensional projections of a z-series through the cell depth, except in some instances specified in figure legends, in which single sections have been used. Images were processed for presentation with Adobe Photoshop 5.0 (Adobe Systems, CA). Standard

electron microscopic techniques for fixation, sectioning and staining or immunodecoration with gold-conjugated antibodies of the samples were used (Nakamura *et al.*, 1995).

## 5 DISCUSSION

### **Kidins220, a new plasma membrane protein enriched in lipid rafts and neurites tips, associates with PKD**

We report here the isolation, molecular cloning and characterisation of a  
10 novel protein interacting with PKD, which presents orthologues in *D. melanogaster*, *C. elegans* and mammals. Several lines of evidence point to a specific functional interaction between PKD and Kidins220: (1) PKD and Kidins220 interact in vitro and in vivo; (2) Kidins220 is directly phosphorylated by PKD, being the only additional phosphorylated protein  
15 in PKD immunoprecipitates (3) Kidins220 contains a consensus motif for PKD-specific phosphorylation and the peptide corresponding to this sequence is strongly phosphorylated by PKD (4) Kidins220 is degraded in vivo when PKD activity is abolished by the use of a dominant negative PKD mutant.

20

The cloning of Kidins220 cDNA reveals no overall similarity to any previously characterised gene products. Kidins220 is an ATP-binding plasma membrane protein highly expressed in brain and PC12 cells, where it accumulates at the neurite tips. Its primary sequence lacks catalytic  
25 domains, but it contains several motifs involved in protein-protein interaction. These include eleven ankyrin repeats, a proline rich region and a carboxy-terminal portion with a high content in serine and threonine residues. Ankyrin repeats have been found to promote cytoskeletal localisation, membrane cytoskeleton bridge formation or protein-protein

interaction (Sedgwick and Smerdon, 1999). Although their role in Kidins220 is presently unknown, these repeats could couple Kidins220 to cytoskeletal components. The association of Kidins220 with the insoluble pellet after Triton X-114 partitioning supports this hypothesis. Interestingly, PKD colocalises with cortactin, an actin binding protein, in invadopodia of epithelial cells (Bowden *et al.*, 1999). Cortactin is enriched in growth cones in neurones (Du *et al.*, 1998), where actin cytoskeleton remodelling takes place (Forscher and Smith, 1988). These findings, together with Kidins220 localisation at the tips of the neurites, suggest a function of PKD/Kidins220 complexes in cytoskeletal reorganisation during neurite extension. In addition to the ankyrin-repeats, Kidins220 has a proline-rich region that may serve as binding site for adaptor modules such as SH3 and WW domains (Nguyen *et al.*, 1998). Finally, Kidins220 carboxy-terminal half is very abundant in serine and threonine (Ser/Thr) residues. This region could participate in interactions with Ser/Thr binding modules (Yaffe and Cantley, 1999).

The association of Kidins220 with the plasma membrane was predicted by the presence of four putative transmembrane domains in its sequence. This has been verified in cell fractionation and immunolocalisation experiments. The plasma membrane distribution of Kidins220 is uneven, presenting higher density in restricted membrane domains. By using two different approaches we have demonstrated the association of Kidins220 to DIGs. Kidins220 is solubilised in non-ionic detergents only when combined with the cholesterol-sequestering agent saponin, indicating that it associates to a cholesterol-rich microenvironment. Double immunolabelling of NGF-treated PC12 cells with Kidins220 and Thy-1, a lipid rafts marker (Madore *et al.*, 1999), showed specific overlapping of Kidins220 and Thy-1 immunofluorescence.

What is the physiological importance of Kidins220 association with DIGs? DIGs have been proposed to function as platforms for the assembly of cytoplasmic and membranous signalling complexes and for polarised membrane trafficking (Brown and London, 1998). Other PKD-interacting proteins, such as BCR, Syk and PLC $\gamma$ 1 (Sidorenko *et al.*, 1996; Cheng *et al.*, 1999) and lipids involved in signal transduction, such as phosphatidylinositol-(4,5)-bisphosphate (PtdIns(4,5)P<sub>2</sub>), have been localised to rafts as well (Hope and Pike, 1996). The selective recruitment of signalling molecules to the plasma membrane is an essential prerequisite for the functioning of various signalling cascades. This is, indeed, the case for PKD, which translocates from the cytosol to the plasma membrane after PDB stimulation (Matthews *et al.*, 1999). In PC12 cells, GFP-PKD migrates to the plasma membrane after phorbol ester stimulation, where Kidins220 is localised (not shown). We suggest a physiological model in which the hydrolysis of DIGs-localised PtdIns(4,5)P<sub>2</sub> by PLCs after receptor stimulation will increase the DAG levels in this lipid microdomains, triggering the recruitment of activated PKD to these areas. Kidins220 would therefore serve as a scaffold for the assembly of a signalling complex at lipid rafts, one of the members being PKD.

**How, where and when does PKD interact and phosphorylate Kidins220?**

Kidins220 and PKD form a stable complex that is not affected by stimulating conditions. Other proteins whose PKD interaction is independent of stimulation have been reported (Sidorenko *et al.*, 1996; Nishikawa *et al.*, 1998; Johannes *et al.*, 1999; Waldron *et al.*, 1999). 14.3.3 proteins are the only example in which phorbol ester stimulation enhances

their association, although they are not a substrate for PKD (Hausser *et al.*, 1999).

PKD is mainly cytosolic in resting PC12 cells. However, it also associates  
5 with membrane fractions, including lipid rafts, which contain Kidins220. The shift of PKD to the plasma membrane after stimulation (Matthews *et al.*, 1999) would increase its transient contact with Kidins220, favouring further phosphorylation of Kidins220 by translocated active PKD molecules. Other authors have found PKD in complexes at the plasma  
10 membrane (Sidorenko *et al.*, 1996; Bowden *et al.*, 1999). In contrast to all these studies, PKD has been reported as a Golgi-localised kinase (Prestle *et al.*, 1996; Jamora *et al.*, 1999). Several PKC isozymes have different localisations depending on the cell type (Goodnight *et al.*, 1995) or undergo a dynamic compartmentalisation (Hug and Sarre, 1993). The  
15 immunolocalisation of PKD to different compartments may therefore be consistent with a model in which PKD plays a variety of roles depending on intracellular localisation, cell type and/or basal metabolic activity.

A critical issue in the understanding of the role of PKD is the identification  
20 of functionally relevant substrates of this kinase. Data from in vitro kinase assays in PC12 cells not only confirm the PKD/Kidins220 interaction, but also show that Kidins220 is phosphorylated by this kinase and that phosphorylation is enhanced under stimulation conditions. Very little is known about the substrate specificity of PKD, considering that there were  
25 not as yet known PKD specific substrates. From previous reports (Van Lint *et al.*, 1995; Dieterich *et al.*, 1996; Nishikawa *et al.*, 1997), PKD appears to prefer substrates with basic residues upstream of the phosphorylatable serine, hydrophobic amino acids (F,L,V) at the +1 position and a critical Leu residue at the -5 position. Kidins220 contains a consensus motif for



PKD phosphorylation at Ser919 (Table I). Results from phosphorylation assays with the synthetic peptide corresponding to this region strongly suggest that this sequence is a target for PKD-Kidins220 phosphorylation in vivo. Although further studies to map the phosphorylated residues are  
5 needed, the data provided here demonstrate that Kidins220 is the first identified physiological substrate for PKD. This conclusion is further supported by the effect of PKD activity on the stability of Kidins220, as discussed below.

#### 10 **PKD regulates Kidins220 trafficking and turnover**

Kidins220 localisation and degradation is modulated by PKD activity, suggesting that PKD participates in the regulation of its substrate's availability. Kinases regulate different substrates by conferring specific phosphorylation patterns that affects conformation, stability, transport and  
15 cellular localisation. The lability of many regulatory proteins is crucially important for their function in the cells. In the case of signalling molecules or plasma membrane receptors, it is apparent that changes in their half-life could have significant impact on the activity of the corresponding signalling cascade. Here we show that PKD activity is a critical determinant of  
20 Kidins220 protein levels and cellular localisation, an indication for their in vivo connected function.

Aspects that remain to be clarified are the molecular mechanisms by which PKD controls Kidins220 levels and localisation. At least two mechanisms  
25 could explain how phosphorylation and dephosphorylation of PKD/Kidins220 complexes might affect Kidins220 stability. The first involves a direct effect of PKD on the pathway controlling the intracellular transport of Kidins220. In this model, the inactivation of PKD activity would impair Kidins220 phosphorylation, which may be required to

- maintain Kidins220 in a functional conformation and/or be essential for the targeting to its final cellular compartment. The final result would be Kidins220 mislocalisation and consequent routing to degradation. The fact that PKD activity has been described as important for protein transport  
5 along the secretory pathway (Prestle *et al.*, 1996; Jamora *et al.*, 1999), strongly supports this being the case. A prediction of this model is that PKD dominant negative mutant would cause the accumulation of Kidins220 in large aggregates, as indeed shown by our experiments.
- 10 Another possible mechanism envisages that Kidins220 dephosphorylation at the plasma membrane induces internalisation and down regulation by a mechanism either lysosomal or proteasome-mediated. Ubiquitination has been implicated as a primary determinant in the control of protein stability. In yeast, ubiquitination participates as an internalisation signal (Hicke,  
15 1997). In mammals, the role of phosphorylation in triggering protein degradation linked to the ubiquitin pathway has been extensively described (Traenckner *et al.*, 1994; Galcheva-Garlova *et al.*, 1995; Mori *et al.*, 1995; Strous *et al.*, 1996; Lu *et al.*, 1998). However, few examples in which phosphorylation is linked to protein stabilisation are available. This is the  
20 case for c-Jun, in which Jun N-terminal kinase (JNK) targets the ubiquitination of non-phosphorylated c-Jun (Fuchs *et al.*, 1996). Similar to JNK, PKD is associated to and phosphorylates Kidins220, suggesting a model in which ubiquitination cause Kidins220 disappearance. This hypothesis correlates well with the presence of a conserved motif for  
25 ubiquitination (Orford *et al.*, 1997) within the Kidins220 sequence at residues 1481-1490 (SKLLPGKKSS) and residues 1637-1445 (SICSEDKKS) (as numbered in Figure 1b). Alternatively, PKD-dependent phosphorylation of Kidins220 could be affecting both secretory and

internalisation pathways, leading to a complex model, which entwines the two mechanisms.

This is the first report presenting evidence for PKD function in protein turnover. Recently, PKD has been implicated in the control of key cellular events such as lymphocyte activation (Sidorenko *et al.*, 1996), metastasis (Bowden *et al.*, 1999), tumour necrosis factor-induced apoptosis (Johannes *et al.*, 1998), growth factor receptor stimulation (Bagowski *et al.*, 1999), Golgi organisation and protein transport (Prestle *et al.*, 1996; Jamora *et al.*, 1999). By introducing this novel PKD role, we propose a unifying mechanism by which PKD may control several cellular functions by altering the traffic and degradation of its interacting substrates. Kidins220, whose cloning and first characterisation is reported herein, is the first PKD-associated protein substrate whose localisation and degradation are affected by PKD activity. The development of specific inhibitors for PKD could therefore be important for the selective degradation of protein substrates that might play a central role in pathogenic processes, like neurodegenerative diseases or cancer.

## **Example 2: Splice variants of the kidins220 protein**

The variants were obtained by direct PCR of cDNA obtained by reverse transcription of total mRNA extracted from PC12 cells. The sequences were verified by re-sequencing both strands twice, and the sequences which differ from the sequence shown in Figure 10 are shown in Figure 8.

PCR products were subcloned in the pGEM-T Easy vector (Promega, Madison, WI) and sequenced by using an ABI Prism 377 DNA sequencer and BigDye Terminator technology (Applied Biosystem, Warrington, UK).

Multiple clones from each RT-PCR or each 5'-RACE reactions ( $> 3$ ) were sequenced in both DNA strands. Variability among different clones was consistently reported in the central portion of Kidins220 cDNA. This variability has been interpreted as possible presence of multiple Kidins220  
5 splicing variants.

Note: variations in the central portion has been verified in full length Kidins220 and are independent from the primer used in RT-PCR. Total RNA was prepared following manufacturer instruction by using the TRIzol  
10 reagent (GIBCO).

**Example 3: Production of antibody specific for Kidins220 phosphorylated at Serine 919**

15 An antibody able to specifically bind to phosphorylated rat Kidins220 but not to unphosphorylated rat Kidins220 was generated by immunising a rabbit with a peptide corresponding to the sequence 908-924 of rat Kidins220 coupled to KLH (Pierce) *via* an additional cysteine at the N-terminus. The antibody is polyclonal and recognises the phosphopeptide  
20 but not the unphosphorylated one as determined by ELISA with immobilised peptide. The entire signal is abolished accordingly by incubation with the phosphorylated but not the unphosphorylated peptide. Affinity purification of the antibody was achieved on a Sulfolink matrix (Pierce) pre-derivatised with the phosphorylated peptide. Elution was with  
25 Glycine 200 mM pH 2.5.

**Example 4: Structural information of Kidins 220****aa17 – 391 Ankyrin repeat region circular profile.**

Ankyrin repeats (ANK) are tandemly repeated modules of about 33 amino acids. They occur in a large number of functionally diverse proteins mainly from eukaryotes. Many ankyrin repeat regions are known to function as protein-protein interaction domains. The conserved fold of the ankyrin repeat unit is known from several crystal and solution structures, e.g. from: - p53-binding protein 53BP2, Cyclin-dependent kinase inhibitor p19Ink4d, NF-kappaB inhibitory protein Ikb-alpha. It has been described as an L-shaped structure consisting of a beta-hairpin and two alpha-helices.

**aa707-748 Domain of unknown function DUF38.**

This domain has no known function. It is found in many proteins from *C. elegans*. The domain is found associated with the F-box. The F-box domain was first described as a sequence motif found in cyclin-F that interacts with the protein SKP1. This motif is present in numerous proteins and serves as a link between a target protein and an ubiquitin-conjugating enzyme. The SCF complex (e.g., Skp1-Cullin-F-box) plays a similar role as an E3 ligase in the ubiquitin protein degradation pathway.

**aa1246-1256 Na<sup>+</sup>/K<sup>+</sup> ATPase C-terminus.**

This extension is specific to the Na<sup>+</sup>/K<sup>+</sup> ATPase subfamily of ATPases. The alpha chains of sodium/potassium-transporting ATPases (H<sup>+</sup>/K<sup>+</sup> and Na<sup>+</sup>/K<sup>+</sup>-ATPase) catalyze the hydrolysis of ATP, coupled with the exchange of sodium and potassium ions across the plasma membrane. The proteins are located in the cell membrane, the ion transport they mediate creating the electro-chemical gradient that provides the energy for the active transport of various nutrients.

**aa1204-1224 PIPLC X DOMAIN:**

Phosphatidylinositol-specific phospholipase X-box domain.  
Phosphatidylinositol-specific phospholipase C, an eukaryotic intracellular  
5 enzyme, plays an important role in signal transduction processes. It  
catalyzes the hydrolysis of 1-phosphatidyl-D-myo-inositol-3,4,5  
triphosphate into the second messenger molecules diacylglycerol and  
inositol- 1,4,5-triphosphate. All eukaryotic PI-PLCs contain two regions of  
homology, sometimes referred to as 'X-box' and 'Y-box'. The order of these  
10 two regions is always the same (NH<sub>2</sub>-X-Y-COOH), but the spacing is  
variable. In most isoforms, the distance between these two regions is only  
50-100 residues but in the gamma isoforms one PH domain, two SH2  
domains, and one SH3 domain are inserted between the two PLC-specific  
domains. The two conserved regions have been shown to be important for  
15 the catalytic activity.

**aa1185-1211 Ly-6 / u-PARA**

A variety of GPI-linked cell-surface glycoproteins are composed of one or  
more copies of a conserved domain of about 100 amino-acid residues.  
20 These proteins are: (i) Urokinase plasminogen activator surface receptor u-  
PAR, which binds to u-PA and is responsible for its proteolysis-  
independent signal transduction activation effects, (ii) The mouse family of  
Ly-6 T-cell antigens (Ly-6A, -6C, -6E, -6F and -6G) and (iii) CD59  
(MACIF or MIRL), an inhibitor of the complement membrane attack  
25 complex action.

**Example 5: Distribution of Kidins220 in developing mice embryos.**

A polyclonal antibody specific for Kidins220 (GSC16) which was raised to the peptide CAASSESTGFGGEERSIL (amino acids 1747 to 1763 of the sequence shown in Figure 1b) was used to investigate the distribution of Kidins220 in developing mice embryos at stage 14, 14 days post coitum. Kidins220 is shown to be expressed in all brain but is particularly concentrated in the tissue underlying the ventricular vesicle and in a layer of migrating neurons which will form the cortex (Figure 20). Kidins220 is present also in the Choroid Plexus, which is differentiating in the pavement of the Fourth ventricle.

**Example 6: Distribution of Kidins220 in the brain.**

Labelling of Kidins220 shows that additional regions of the brain that contain high levels of Kidins220 are the dorsal root ganglia, the trigeminal ganglion and the differentiating neuroepithelium of the inner ear (cochlear canal) (Figure 21). Other structures that appear labelled are the primordia of the tactile and sinus air follicle and the developing external ear.

**Example 7: SAGE analysis of human Kidins220 expression.**

The human Kidins220 nucleotide sequence was analysed using the SAGEmap vNorthern tool provided by the Cancer Genome Anatomy Project (<http://cgap.nci.nih.gov/>). The sequence is shown to correspond to the Unigene cluster Hs.9873 with a corresponding SAGE tag "CTAAAAAAAAA". *In silico* analysis of the relative number of tags encountered in various libraries (which number is an indication of the relative levels of expression of the gene in the corresponding tissue;

[www.sagenet.org](http://www.sagenet.org)) reveals that Kidins220 is over-expressed in a variety of tumours, especially in glioblastoma multiforme and prostate cancers (Figure 22).

5    **Example 8: Nuclear localisation of Kidin220 in the astrocytoma-glioblastoma U87MG cell line.**

Analysis of Kidins220 expression in a human cell line derived from an astrocytoma/glioblastoma (U87MG, ICRF cell depository) confirmed the  
10    high level of expression of this protein in this type of neoplasm of the CNS. This high level of expression is accompanied by the shift of Kidins220 localisation from the plasma membrane to the cell nucleus (Figure 23). Although the molecular mechanism responsible for this differential trafficking remains unknown, the nuclear localisation of Kidins220 could  
15    represent a diagnostic sign to follow neoplastic progression and to distinguish neoplastic glia-derived cells from neighbouring neurons.

**REFERENCES FOR EXAMPLE 1**

- Bagowski, C. P., Stein-Gerlach, M., Choidas, A., and Ullrich, A. (1999).  
20    Cell-type specific phosphorylation of threonines T654 and T669 by PKD defines the signal capacity of the EGF receptor. *EMBO J.* 18, 5567-5576.
- Bordier, C. (1981). Phase separation of integral membrane proteins in Triton X-114 solution. *J. Biol. Chem.* 256, 1604-1607.
- Bowden, E. T., Barth, M., Thomas, D., Glazer, R. I., and Mueller, S. C.  
25    (1999). An invasion-related complex of cortactin, paxillin and PKC $\mu$  associates with invadopodia at sites of extracellular matrix degradation. *Oncogene* 18, 4440-4449.
- Brown, D., and London, E. (1998). Function of lipid rafts in biological membranes. *Annu Rev Cell Dev. Biol.* 14, 111-136.



- Bukovsky, A., Presl, J., Zidovsky, J., and Mancal, P. (1983). The localization of Thy-1.1, MRC OX 2 and Ia antigens in the rat ovary and fallopian tube. *Immunology* 48, 587-96.
- Cheng, P. C., Dykstra, M. L., Mitchell, R. N., and Pierce, S. K. (1999). A  
5 role of lipid rafts in B cell antigen receptor signaling and antigen targeting. *J. Exp. Med.* 190, 1549-1560.
- Dieterich, S., Herget, T., Link, G., Bottinger, H., Pfizenmaier, K., and Johannes, F. J. (1996). In vitro activation and substrates of recombinant baculovirus expressed human protein kinase C  $\mu$ . *FEBS Lett.* 381, 183-187.
- 10 Dittie, A., Thomas, L., Thomas, G., and Tooze, S. (1997). Interaction of furin in immature secretory granules from neuroendocrine cells with the AP-1 adaptor complex is modulated by casein kinase II phosphorylation. *EMBO J.* 16, 4859-4870.
- Divecha, N., and Irvine, R. F. (1995). Phospholipid signaling. *Cell* 80, 269-  
15 278.
- Du, Y., Weed, S. A., Xiong, W.-C., Marshall, T. D., and Parsons, J. T. (1998). Identification of a novel cortactin SH3 domain-binding protein and its localisation to growth cones of cultured neurones. *Mol. Cell. Biol.* 18, 5838-5851.
- 20 Exton, J. H. (1990). Signaling through phosphatidylcholine breakdown. *J. Biol. Chem.* 265, 1-4.
- Forscher, P., and Smith, S. J. (1988). Actions of cytochalasins on the organisation of actin filaments and microtubules in a neuronal growth cone. *J. Cell Biol.* 107, 1505-1516.
- 25 Fuchs, S. Y., Dolan, L., Davis, R. J., and Ronai, Z. (1996). Phosphorylation-dependent targeting of c-Jun ubiquitination by Jun N-kinase. *Oncogene* 13, 1531-1535.

- Galcheva-Garlova, Z., Theroux, S. J., and Davis, R. J. (1995). The epidermal growth factor is covalently linked to ubiquitin. *Oncogene 11*, 2649-2655.
- Goodnight, J., Mischak, H., Kolch, W., and Mushinski, J. F. (1995).  
5 Immunocytochemical localisation of eight protein kinase C isotypes overexpressed in NIH 3T3 fibroblasts. *J. Biol. Chem. 270*, 9991-10001.
- Harder, T., and Simons, K. (1997). Caveolae, DIGs, and the dynamics of sphingolipid-cholesterol microdomains. *Curr. Opin. Cell Biol. 9*, 534-42.
- Hausser, A., Storz, P., Link, G., Stoll, H., Liu, Y.-C., Altman, A.,  
10 Pfizenmaier, K., and Johannes, F. (1999). Protein Kinase C  $\mu$  Is Negatively Regulated by 14-3-3 Signal Transduction Proteins. *J. Biol. Chem. 274*, 9258-9264.
- Herschmann, H. R. (1991). Primary response genes induced by growth factors and tumor promoters. *Annu. Rev. Biochem. 60*, 281-319.
- 15 Hicke, L. (1997). Ubiquitin-dependent internalisation and down-regulation of plasma membrane proteins. *FASEB J. 11*, 1215-1226.
- Hope, H. R., and Pike, L. J. (1996). Phosphoinositides and phosphoinositide-utilising enzymes in detergent insoluble lipid domains. *Mol. Biol. Cell 7*, 843-851.
- 20 Hoss, M., Robins, P., Naven, T. J., Pappin, D. J., Sgouros, J., and Lindahl, T. (1999). A human DNA editing enzyme homologous to the Escherichia coli DnaQ/MutD protein. *EMBO J. 18*, 3868-75.
- Hug, H., and Sarre, T. F. (1993). Protein Kinase C isoenzymes: divergence in signal transduction. *Biochem. J. 291*, 329-343.
- 25 Iglesias, T., Matthews, S., and Rozengurt, E. (1998a). Dissimilar phorbol ester binding properties of the individual cystein rich motifs of protein kinase D. *FEBS Lett. 437*, 19-23.

- Iglesias, T., and Rozengurt, E. (1998b). Protein kinase D activation by mutations within its pleckstrin homology domain. *J. Biol. Chem.* 273, 410-416.
- Iglesias, T., Waldron, R., and Rozengurt, E. (1998c). Identification of in vivo phosphorylation sites required for protein kinase D activation. *J. Biol. Chem.* 273, 27662-27667.
- Ilangumaran, S., and Hoessli, D. (1998). Effects of cholesterol depletion by cyclodextrin on the sphingolipid microdomains of the plasma membrane. *Biochem. J.* 335, 433-440.
- 10 Jamora, C., Yamanouye, N., Lint, J. V., Vandenheede, J. R., Faulkner, D. J., and Malhotra, V. (1999). Gbg-mediated regulation of Golgi organisation is through the direct activation of Protein Kinase D. *Cell* 98, 59-68.
- Johannes, F. J., Prestle, J., Eis, S., Oberhagemann, P., and Pfinzenmaier, K. (1994). PKC $\mu$  is a novel, atypical member of the protein kinase C family. *J. Biol. Chem.* 269, 6140-6148.
- 15 Johannes, F. J., Horn, J., Link, G., Haas, E., Siemienski, K., Wajant, H., and Pfinzenmaier, K. (1998). Protein Kinase C $\mu$  downregulation of tumor-necrosis-factor-induced apoptosis correlates with enhanced expression of nuclear-factor-kB-dependent protective genes. *Eur. J. Biochem.* 257, 47-54.
- 20 Johannes, F. J., Hausser, A., Storz, P., Truckenmuller, Link, G., Kawakami, T., and Pfinzenmaier, K. (1999). Bruton's tyrosine kinase (Btk) associates with protein kinase C $\mu$ . *FEBS Lett.* 462, 68-72.
- Kikkawa, U., and Nishizuka, Y. (1986). The role of protein kinase C in transmembrane signalling. *Ann. Rev. Cell Biol.* 2, 149-154.
- 25 Kozak, M. (1992). Regulation of translation in eucaryotic systems. *Annu. Rev. Cell Biol.* 8, 197-225.
- Linstedt, A., and Hauri, H. (1993). Giantin, a novel conserved Golgi membrane protein containing a cytoplasmic domain of at least 350 kDa. *Mol. Biol. Cell* 4, 679-693.

- Lu, Z., Liu, D., Hornia, A., Devonish, W., Pagano, M., and Foster, D. (1998). Activation of protein kinase C triggers its ubiquitination and degradation. *Mol. Cell Biol.* 18, 839-845.
- Madore, N., Smith, K. L., Graham, C. H., Jen, A., Brady, K., Hall, S., and  
5 Morris, R. (1999). Functionally different GPI proteins are organised in different domains on the neuronal surface. *EMBO J.* 18, 6917-6926.
- Marshall, C. (1995). Specificity of receptor tyrosine kinase signaling: transient versus sustained extracellular signal-regulated kinase activation. *Cell* 80, 179-185.
- 10 Matthews, S., Iglesias, T., Cantrell, D., and Rozengurt, E. (1999). Dynamic re-distribution of protein kinase D (PKD) as revealed by a GFP-PKD fusion protein: dissociation from PKD activation. *FEBS Lett.* 457, 515-521.
- Mellor, H., and Parker, P. J. (1998). The extended protein kinase C superfamily. *Biochem. J.* 332, 281-292.
- 15 Mori, S., Tanaka, K., Omura, S., and Saito, Y. (1995). Degradation process of ligand-stimulated platelet-derived growth factor B-receptor involves ubiquitin-proteasome proteolytic pathway. *J. Biol. Chem.* 270, 29447-29452.
- Nagase, T., Ishikawa, K., Kikuno, R., Hirosawa, M., Nomura, N., and  
20 Ohara, O. (1999). Prediction of the coding sequences of unidentified human genes. XV. The complete sequences of 100 new cDNA clones from brain which code for large proteins *in vitro*. *DNA Res.* 6, 337-345.
- Nakamura, N., Rabouille, C., Watson, R., Nilsson, T., Hui, N., Shusarewicz, P., Kreis, T. E., and Warren, G. (1995). Characterization of a cis-Golgi  
25 matrix protein, GM130. *J. Cell Biol.* 131, 1715-1726.
- Newton, A. (1997). Regulation of protein kinase C. *Curr. Opin. Cell Biol.* 9, 161-167.

- Nguyen, J. T., Turck, C. W., Cohen, F. E., Zuckerman, R. N., and Lim, W. A. (1998). Exploiting the basis of proline recognition by SH3 and WW domains: desing of N-substituted inhibitors. *Science* 282, 2088-2092.
- Nishikawa, K., Toker, A., Johannes, F., Songyang, Z., and Cantley, L. 5 (1997). Determination of the specific substrate sequence motifs of protein kinase C isozymes. *J. Biol. Chem.* 272, 952-960.
- Nishikawa, K., Toker, A., Wong, K., Marignani, P., Johannes, F., and Cantley, L. (1998). Association of protein kinase C m with type II phosphatidylinositol 4-Kinase and type I phosphatidylinositol-4-phosphate 10 5-kinase. *J. Biol. Chem.* 273, 23126-33.
- Nishizuka, Y. (1992). Intracellular signaling by hydrolysis of phospholipids and activation of protein kinase C. *Science* 258, 607-614.
- Orford, K., Crockett, C., Jensen, J. P., Weissman, A. M., and Byers, S. W. (1997). Serine phosphorylation-regulated ubiquitination and degradation of 15 b-catenin. *J. Biol. Chem.*
- Prestle, J., Pfizenmaier, K., Brenner, J., and Johannes, F. J. (1996). Protein Kinase  $\mu$  is located at the Golgi compartment. *J. Cell Biol.* 134, 1401-1410.
- Ron, D., and Kazanietz, M. G. (1999). New insights into the regulation of protein kinase C and novel phorbol ester receptors. *FASEB J.* 13, 1658- 20 1676.
- Rozengurt, E. (1986). Early signals in the mitogenic response. *Science* 234, 161-166.
- Rozengurt, E., Sinnett-Smith, J., Lint, J. V., and Valverde, A. M. (1995). Protein kinase D (PKD): A novel target for diacylglycerol and phorbol 25 esters. *Mut. Res.* 333, 153-160.
- Sedgwick, S. G., and Smerdon, S. J. (1999). The ankyrin repeat: a diversity of interactions on a common structural framework. *Trends Biol. Sci.* 24, 311-316.

- Sidorenko, S., Law, C., Klaus, S., Chandran, K., Takata, M., Kurosaki, T., and Clark, E. (1996). Protein kinase C mu (PKC  $\mu$ ) associates with the B antigen receptor complex and regulates lymphocyte signaling. *Immunity* 5, 353-363.
- 5 Soldati, T., and Perriard, J.-C. (1991). Intracompartimental sorting of essential myosin light chains: molecular dissection and in vivo monitoring by epitope tagging. *Cell* 66, 277-289.
- Strous, G., Kerkhof, P. V., Govers, R., Ciechanover, A., and Schwartz, A. L. (1996). The ubiquitin conjugation system is required for ligand-induced  
10 endocytosis and degradation of the growth hormone receptor. *EMBO J.* 15, 3806-3812.
- Traenckner, E. B. M., Wilk, S., and Baeuerle, P. A. (1994). A proteasome inhibitor prevents activation of NF-kB and stabilises a newly phosphorylated form of I $\kappa$ B-s that is still bound to NF-kB. *EMBO J.* 13,  
15 5433-5441.
- Valverde, A. M., Sinnott-Smith, J., Van Lint, J., and Rozengurt, E. (1994). Molecular cloning and characterization of protein kinase D: A Target for diacylglycerol and phorbol esters with a distinctive catalytic domain. *Proc. Natl. Acad. Sci. USA* 91, 8572-8576.
- 20 Van Lint, J., Sinnott-Smith, J., and Rozengurt, E. (1995). Expression and characterization of PKD, a phorbol ester and diacylglycerol-stimulated serine protein kinase. *J. Biol. Chem.* 270, 1455-1461.
- Waldron, R., Iglesias, T., and Rozengurt, E. (1999). Phosphorylation dependent PKD activation. *Electrophoresis* 20, 382-390.
- 25 Waldron, R., Iglesias, T., and Rozengurt, E. (1999). The pleckstrin homology domain of protein kinase D interacts preferentially with the eta isoform of protein kinase C. *J. Biol. Chem.* 274, 9224-9230.
- Wooten, M., Zhou, G., Wooten, M., and Seibenhener, M. (1997). Transport of protein kinase C isoforms to the nucleus of PC12 cells by nerve growth

- factor: Association of atypical zeta-PKC with the nuclear matrix. *J. Neurosci. Res.* 49, 393-403.
- Yaffe, M. B., and Cantley, L. C. (1999). Signal transduction. Grabbing phosphoproteins. *Nature* 402, 30-31.
- 5 Zugaza, J. L., Sinnott-Smith, J., Lint, J. V., and Rozengurt, E. (1996). Protein kinase D (PKD) activation in intact cells through a protein kinase C-dependent signal transduction pathway. *EMBO J.* 15, 6220-6230.
- Zugaza, J. L., Waldron, R. T., Sinnott-Smith, J., and Rozengurt, E. (1997). Bombesin, vasopresin, endothelin, bradykinin, and platelet derived growth  
10 factor rapidly activate protein kinase D through a protein kinase C-dependent signal transduction pathway. *J. Biol. Chem.* 272, 23952-23960.

CLAIMS

1. A polynucleotide comprising a nucleotide sequence which encodes the polypeptide whose amino acid sequence is given in Figure 1b or Figure 11 or a fragment or variant or fusion thereof, or a fusion of said fragment or variant, or a complement thereof, provided that the polynucleotide is not any one of the clones corresponding to the GenBank Accession Nos AL133620, AB033076 AC007175, F36H1, or any Accession No. given in Figure 14a or 14b.
2. A polynucleotide according to Claim 1 comprising a nucleotide sequence encoding a polypeptide whose amino acid sequence is given in Figure 8.
3. A polynucleotide according to Claim 1 comprising the nucleic acid sequence given in Figure 7.
4. A polynucleotide according to Claim 1 comprising a nucleic acid sequence given in Figure 8.
5. A polynucleotide according to Claim 1 which encodes a polypeptide with at least 65% identity to the amino acid sequence given in Figure 1b or Figure 11.
6. A polynucleotide according to Claim 5 which encodes a polypeptide with at least 85% identity to the amino acid sequence given in Figure 1b or Figure 11.



7. A polynucleotide according to Claim 5 which encodes a polypeptide with at with at least 95% identity to the amino acid sequence given in Figure 1b or Figure 11.
- 5 8. A polynucleotide according to any one of Claims 1 to 7 which encodes any one of the peptides CAASSESTGFGGEERSIL, RQMQRITTRQMSFDLTK or RQMQRITTRQMAFDLTK or a part thereof which is no less than 10 contiguous amino acids.
- 10 9. A polynucleotide according to any one of the preceding claims which is at least 30 nucleotides in length.
10. A polynucleotide according to Claims 1 to 9 which contains no introns.
- 15 11. A polynucleotide which is no more than 250bp and which is able to selectively hybridise to any one of the polynucleotides as defined in Claims 1 to 10.
- 20 12. A polynucleotide according to Claim 11 which is between 10 and 200 nucleotides in length.
13. A polynucleotide according to any one of the preceding claims further comprising a detectable label.
- 25 14. An expression vector comprising a polynucleotide according to any one of Claims 1 to 10.

15. An expression vector according to Claim 14 which is capable of expressing a polypeptide whose amino acid sequence is shown in Figure 1b or Figure 11, or a fragment or variant or fusion thereof or a fusion of said fragment or variant.
- 5
16. A replicable vector comprising a polynucleotide as defined in any one of Claims 1 to 10.
- 10
17. A polynucleotide comprising the sequence shown in Figure 16 or a fragment or variant thereof provided that it is not the clone corresponding to GenBank Accession No AC012495.
- 15
18. A polynucleotide according to Claim 17 which is capable of promoting transcription.
19. A polynucleotide according to Claim 18 wherein the transcription is neuron specific.
- 20
20. A polynucleotide according to any one of Claims 17 to 19 operably linked to a polynucleotide encoding a polypeptide.
21. A polynucleotide according to Claim 20 wherein the polypeptide is a therapeutic polypeptide.
- 25
22. A polynucleotide according to Claim 20 wherein the polynucleotide encoding a polypeptide is a polynucleotide according to any one of Claims 1 to 10, or a polypeptide capable of expressing a polypeptide whose amino acid sequence is given in Figure 1b or Figure 11 or a

fragment or variant or fusion thereof, or a fusion of said fragment or variant.

23. A polynucleotide according to Claim 20 wherein the polypeptide is a  
5 detectable polypeptide.
24. A polynucleotide according to Claim 20 wherein the polypeptide is Protein Kinase D or a functional mutant thereof.
- 10 25. A nucleic acid vector comprising a polynucleotide as defined in any one of Claims 17-24.
26. An expression vector according to either of Claims 14 or 15 or a  
15 polynucleotide according to any one of Claims 17 to 25 for use in therapy.
27. A recombinant host cell containing a polynucleotide according to any  
one of Claims 1 to 10 or Claims 17-24 or a replicable vector or an  
expression vector according to any one of Claims 14 to 16 or Claim  
20 25.
28. A host cell according to Claim 27 wherein the host cell further  
comprises a polynucleotide capable of expressing Protein Kinase D  
or a functional equivalent thereof.
- 25 29. A recombinant host cell according to Claims 26 to 28 where the host  
cell is a bacterial cell.

30. A recombinant host cell according to Claims 26 to 28 where the host cell is a mammalian cell.
31. A recombinant host cell according to Claims 26 to 28 where the host  
5 cell is an insect cell.
32. Use of a polynucleotide according to any one of Claims 1 to 10 and 14 to 16 in the production of the encoded polypeptide.
- 10 33. A polypeptide comprising the amino acid sequence given in Figure 1b or Figure 11 or a fragment or variant or fusion thereof or a fusion of said fragment or variant provided that said polypeptide is not the polypeptide encoded by GenBank Accession No AB033076.
- 15 34. A polypeptide according to Claim 33 which has an amino acid sequence with at least 65% identity to the amino acid sequence of Figure 1b or Figure 11.
- 20 35. A polypeptide according to Claim 33 which has an amino acid sequence with at least 85% identity to the amino acid sequence of Figure 1b or Figure 11.
- 25 36. A polypeptide according to Claim 33 which has an amino acid sequence with at least 95% identity to the amino acid sequence of Figure 1b or Figure 11.
37. A polypeptide according to Claim 33 comprising an amino acid sequence given in Figure 1b wherein the amino acids from residues

1140 to 1232 are replaced with the amino acids whose sequence is shown between the square brackets in Figure 8.

- 5 38. A polypeptide according to any one of Claims 33 to 37 wherein the serine located at position 919 as numbered in Figure 1b, or an equivalent position, is phosphorylated.
- 10 39. A polypeptide according to any one of Claims 33 to 37 wherein the serine located at position 919 as numbered in Figure 1b, or an equivalent position, is altered to an amino acid other than serine.
- 15 40. A polypeptide according to any one of Claims 33 to 37 wherein the tyrosine residue located at any one of positions 409 or 1453 as numbered in Figure 1b, or an equivalent position, is altered to an amino acid other than tyrosine.
- 20 41. A polypeptide according to any one of Claims 33 to 37 wherein any one of the lysine residues located at any one of the positions 473, 1482, 1487, 1488, 1643 or 1644 as numbered in Figure 1b, or an equivalent position, is altered to an amino acid other than lysine.
- 25 42. A polypeptide comprising the sequence (I/L)XRQM(S/A)J wherein X is any amino acid, J is a hydrophobic amino acid, (I/L) represents isoleucine or leucine and (S/A) represents serine or alanine and which is not the polypeptide encoded by the nucleotide sequence in GenBank Accession No AB033076 and which is not  $\mu$  peptide (AALVRQMSVAFFFK).

43. A polypeptide comprising the amino acid sequence CAASSESTGFGGEERSIL or a variant or fragment or fusion thereof or a fusion of said fragment or variant or an analogue thereof provided that it is not the polypeptide encoded by the nucleotide  
5 sequence in GenBank Accession No AB033076.
44. A polynucleotide which encodes a polypeptide according to Claims 33 to 43.
- 10 45. A method of producing a peptide or polypeptide according to any one of Claims 33 to 43 comprising expressing a polynucleotide of any one of Claims 1 to 10 which encodes said peptide or polypeptide.
- 15 46. A method of producing a peptide or polypeptide according to any one of Claims 33 to 43 comprising culturing a host cell according to any one of Claims 27 to 31.
- 20 47. A method according to Claim 45 wherein the polynucleotide is translated in an *in vitro* translation system.
48. A method according to Claims 45 to 47 wherein the method comprises the further step of isolating said polypeptide.
- 25 49. A method according to Claim 48 wherein the isolation step employs an antibody which selectively binds the expressed polypeptide.
50. An antibody which specifically binds to a polypeptide whose amino acid sequence is given in Figure 1b or Figure 11, or a fragment or variant of the said polypeptide.

51. An antibody according to Claim 50 which does not bind to a variant wherein the amino acids from residues 1140 to 1232 in Figure 1b are replaced with the amino acid whose sequence is shown between the square brackets in Figure 8.
52. An antibody according to Claim 50 which specifically binds to a variant wherein the amino acids from residues 1140 to 1232 in Figure 1b are replaced with the amino acid whose sequence is shown between the square brackets in Figure 8 and which antibody does not bind to a polypeptide whose amino acid sequence is given in Figure 1b.
53. An antibody which specifically binds to a polypeptide comprising any of the amino acid sequences TEDAAEGLPS, CHRQLTVTEF, RQVQKLQAAV, DQNNGLAAVP, HPFYNRAAVP, HPFYNRANIN and FFAPYLYTPR.
54. An antibody according to Claim 50 whose binding can distinguish between one or a selection made from the polypeptide whose amino acid sequence is given in Figure 1b and the polypeptide variants wherein the amino acids from residues 1140 to 1232 in Figure 1b are replaced with the amino acid whose sequence is shown between the square brackets in Figure 8.
55. An antibody according to Claim 50 which specifically binds to the amino acid sequence CAASSESTGFGEERSIL.

56. An antibody according to Claim 50 which specifically binds to a polypeptide whose amino acid sequence comprises at least five consecutive residues including residue 919 of the polypeptide as numbered in Figure 1b or an equivalent position of the polypeptide of Figure 11.
57. An antibody according to Claim 56 which specifically binds to the amino acid sequence RQMQRITRQMSFDLTK.
58. An antibody according to any one of Claims 50, 51, 56 and 57 wherein the serine at residue 919 as numbered in Figure 1b, or an equivalent residue in Figure 11, is phosphorylated.
59. An antibody according to any one of Claims 50, 51, 56 and 57 wherein the serine at residue 919 as numbered in Figure 1b, or an equivalent residue in Figure 11, is non-phosphorylated.
60. An antibody as defined in any one of Claims 50 to 59 which is a monoclonal antibody.
61. An antibody according to any one of Claims 50 to 59 further comprising a detectable label.
62. A method of producing an antibody according to any one of Claims 50 to 59.
63. A method of modulating the activity of a polypeptide according to any one of Claims 33 to 42 the method comprising contacting the



polypeptide with an antibody according to any one of Claims 50 to 61.

- 5 64. A method according to Claim 63 wherein addition of an antibody according to any one of Claims 50, 51, 56, 57 and 59 prevents a region of the polypeptide from interacting with another protein.
- 10 65. A method according to Claim 63 wherein addition of an antibody according to any one of Claims 50, 51, 56, 57 and 59 prevents phosphorylation of serine 919 of the polypeptide as numbered in Figure 1b or an equivalent residue.
- 15 66. A method of modulating the function of a polypeptide according to any one of Claims 33 to 37 when in the presence of Protein Kinase D, the method comprising inhibiting the phosphorylating ability of Protein Kinase D.
- 20 67. A method of modulating the function of the polypeptide according to any one of Claims 33 to 42 comprising reducing the ability of said polypeptide to bind ATP.
- 25 68. A method of identifying an agent which modulates the promoter activity of the polynucleotide according to any one of Claims 17-24, said method comprising
- i) introducing said polynucleotide into a suitable cell;
  - ii) exposing a test agent to said cell;
  - iii) detecting whether (and, optionally, to what extent) the promoter activity of the polynucleotide has been altered.

69. A method of identifying an agent which modulates the activity of the polypeptide according to Claims 33 to 42, said method comprising:

- 5           i) exposing a test agent to said polypeptide and;  
          ii) detecting whether (and, optionally, to what extent) the activity of said polypeptide has been altered.

70. A method of identifying an agent which modulates the activity of Protein Kinase D, the method comprising:

- 10           i) exposing Protein Kinase D or a functional equivalent thereof to a peptide or polypeptide comprising the amino acid sequence (I/L)XRQMSJ, where X is any amino acid, J represents a hydrophobic residue and (I/L) represents isoleucine or leucine or to a peptide or polypeptide comprising at least five consecutive  
15           residues including residue 919 of the polypeptide as numbered in Figure 1b; and  
          ii) detecting, in the presence or absence of the test agent whether (and, optionally, to what extent) said peptide or polypeptide has been phosphorylated.

20

71. A method according to Claim 70 wherein the test agent is exposed to Protein Kinase D or a functional equivalent thereof prior to the Protein Kinase D or a functional equivalent thereof being exposed to the said peptide or polypeptide.

25

72. A method according to Claim 70 or 71 wherein the polypeptide comprising the amino acid sequence (I/L)XRQMSJ, has an amino acid sequence with at least 65% identity to the amino acid sequence of Figure 1b.

73. A method of identifying an agent which modulates the interaction between Protein Kinase D and the polypeptide whose amino acid sequence is shown in Figure 1b or a fragment or variant thereof or fusion or a fusion of said fragment or variant, the method comprising detecting in the presence or absence of the test agent whether Protein Kinase D interacts with the said polypeptide.
74. An agent identifiable by the method of any of Claims 68 to 72.
75. An agent identified by the method of any of Claims 68 to 72.
76. An agent identifiable or identified by the method of any of Claims 68 to 72 for use in medicine.
77. A polynucleotide according to any one of Claims 1 to 12 or Claims 17 to 24 or a polypeptide according to any one of Claims 33 to 43 for use in medicine.
78. A polynucleotide according to Claims 1 to 12 or Claims 17 to 24 further comprising a vector suitable for use in gene therapy.
79. A pharmaceutical composition comprising an agent identifiable or identified by the method of any of Claims 68 to 72 or a polynucleotide according to any one of Claims 1 to 12 or a polypeptide according to any one of Claims 33 to 43 and a pharmaceutically acceptable carrier.

80. Use of an agent identifiable or identified by the method of any of Claims 64 to 67 or a polynucleotide according to any one of Claims 1 to 12 or Claims 17 to 24 or a polypeptide according to any one of Claims 33 to 43 in the manufacture of a medicament for treatment of neurodegenerative disease.
81. Use of an agent identifiable or identified by the method of any of Claims 64 to 67 or a polynucleotide according to any one of Claims 1 to 12 or Claims 17 to 24 or a polypeptide according to any one of Claims 33 to 43 for treating neurodegenerative disease.
82. A method of modulating the function of Protein Kinase D the method comprising the step of contacting Protein Kinase D with a peptide or polypeptide according to any one of Claims 33 to 39 and 42.
83. A kit of parts comprising Protein Kinase D or a functional equivalent thereof and a polypeptide whose amino acid sequence is shown in Figure 1b as a fragment or variant or fusion thereof or a polypeptide according to Claim 42.
84. A kit of parts according to Claim 83 wherein the Protein Kinase D is a constitutively active mutant.
85. A kit of parts according to Claim 83 or 84 further comprising a means for detecting phosphorylation of a protein.
86. A mutant animal wherein if the organism contains in its genome a polynucleotide which encodes a polypeptide comprising the amino

acid sequence given in Figure 1b or a fragment or a variant thereof the polynucleotide in the mutant organism is mutated or absent.

87. A mutant organism according to Claim 86 which is a mouse.

88. A polypeptide comprising any one of the amino acid sequences  
TEDAAEGLPS, CHRQLTVTEF, RQVQKLQAAV,  
DQNNGLAAVP, HPFYNRAAVP and HPFYNRANIN.

89. A polypeptide according to Claim 88 which has between 10 and 50 amino acids.

90. A method of diagnosing cancer in a patient comprising the steps of  
(i) determining the level of expression of the Kidins220 gene in a sample from the patient containing protein or nucleic acid;  
(ii) comparing the level of expression of step (i) with that in a control sample.

91. A method of monitoring the progression of cancer in a patient comprising the steps of  
(i) determining the level of expression of the Kidins220 gene in a sample from the patient containing protein or nucleic acid;;  
(ii) comparing the level of expression of step (i) with that in a previous sample from the same patient.

92. A method according to Claims 90 or 91 wherein the level of expression of said gene is determined using an antibody capable of selectively binding to the Kidins220 polypeptide or a polynucleotide according to any one of Claims 1 to 13.

93. A method according to Claim 92 wherein the antibody is one according to any one of Claims 50 to 61.

5 94. A method of diagnosing cancer in a patient comprising the steps of  
(i) determining the phosphorylation status of the Kidins220 polypeptide in a sample from the patient containing protein;  
(ii) comparing the phosphorylation status of step (i) with that in a control sample.

10

95. A method of diagnosing cancer in a patient comprising the steps of  
(i) determining the cellular location of the Kidins220 polypeptide in a sample from the patient containing cells;  
(ii) comparing the cellular location of step (i) with that in a control sample.

15

96. A method of monitoring the progression of cancer in a patient comprising the steps of  
(i) determining the phosphorylation status of the Kidins220 polypeptide in a sample from the patient containing protein;  
(ii) comparing the phosphorylation status of step (i) with that in a control sample.

20

97. A method of monitoring the progression of cancer in a patient comprising the steps of  
(i) determining the cellular location of the Kidins220 polypeptide in a sample from the patient containing cells;  
(ii) comparing the cellular location of step (i) with that in a control sample.

25

98. A method of treating cancer in a patient comprising administering to the patient an agent which inhibits function of the Kidins220 gene or its product.

5

99. A method according to Claim 98 comprising administering to the patient a polynucleotide which is antisense to a polynucleotide according to any one of Claims 1 to 10 or 17 to 20 or 22, or an antibody according to any one of Claims 50 to 61.

10

100. A method according to any one of Claims 90 to 99 wherein the cancer is glioblastoma multiforme or prostate cancer.

15

101. An agent capable of inhibiting the function of the Kidins220 gene or its product, an antibody according to any one of Claims 51 to 61 or a polynucleotide which is antisense to a polynucleotide according to any one of Claims 1 to 10 or 17 to 20 or 22, for use in medicine.

20

102. Use of an agent capable of inhibiting the function of the Kidins220 gene or its product in the manufacture of a medicament for treating cancer.

25

103. Use according to Claim 102 wherein the agent is an antibody according to any one of Claims 51 to 61 or a polynucleotide which is antisense to a polynucleotide according to any one of Claims 1 to 10 or 17 to 20 or 22.

104. Use of an agent which is capable of detecting the expression of the Kidins220 gene in the manufacture of a diagnostic reagent for

diagnosing or prognosing cancer or for monitoring the progression of cancer in a patient.

105. Use according to Claim 104 wherein the agent is an antibody  
5 according to any one of Claims 50 to 61 or a polynucleotide according to any one of Claims 1 to 13.
106. Use according to any one of Claims 102 to 105 wherein the cancer is glioblastoma multiforme or prostate cancer.
107. A compound comprising (i) a moiety which selectively binds the  
10 Kidins220 polypeptide or a variant thereof and (ii) a further moiety.
108. A compound according to Claim 107 wherein the moiety which  
15 selectively binds is an antibody.
109. A compound according to either of Claims 107 or 108 wherein the further moiety is a readily detectable moiety.
110. A compound according to any one of Claims 107 or 108 wherein the  
20 further moiety is a directly or indirectly cytotoxic moiety.
111. A nucleic acid molecule encoding a compound according to Claim  
107.
112. A method of imaging cancer in the body of an individual the method  
25 comprising administering to the individual an effective amount of a compound according to Claim 109.



113. A method of diagnosing or prognosing cancer in an individual the method comprising administering to the individual an effective amount of a compound according to Claim 109.
- 5 114. A method of treating an individual in need of treatment, the method comprising administering to the individual an effective amount of a compound according to any one of Claims 107, 108 or 110.
- 10 115. A pharmaceutical composition comprising a compound according to any one of Claims 107 to 110 and a pharmaceutically acceptable carrier.
116. A compound according to any one of Claims 107 to 110 for use in medicine.
- 15 117. Use of a compound according to Claim 109 in the manufacture of an agent for imaging cancer in a body of an individual.
- 20 118. Use of a compound according to Claim 109 in the manufacture of a diagnostic or prognostic agent for cancer.
119. Use of a compound according to any one of Claims 107, 108 or 110 in the manufacture of a medicament for treating cancer.
- 25 120. A kit of parts comprising a compound according to Claim 107 and any one of a directly or indirectly cytotoxic moiety or a readily detectable moiety to which the said compound is able to bind *via* its further moiety.

Figure 1 (page 1 of 2)

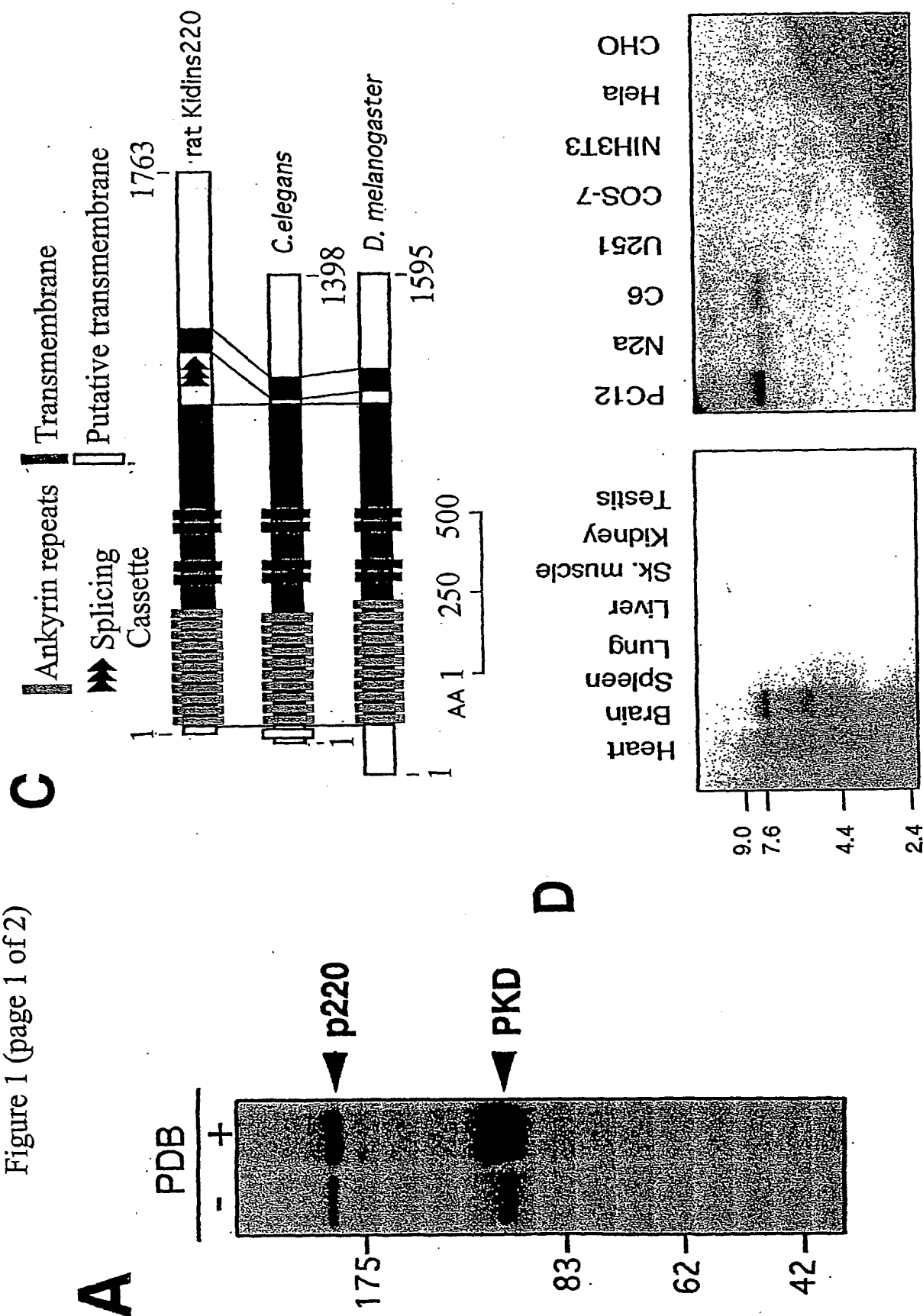


Figure 1 (page 2 of 2)

B

1 MSVLISQSVINYVEEENIPALKALLEKCKDVDERNECGQTPLMLAAEQGNVEIVKELLKNGANCNLEDL DNWTALISASKEGHIHIVEEL  
 91 LKSGASLEHRD MGGWTALMWACYKGRDQVVELLSHGANSPTGL QYSVYPIIWAAGRGHADIVHLLLLQNGAKVNCSDK YGTTPLVWAAR  
 181 KGHLECVKHLAMGADVDEGA NSTTALIVAVKGGYTQSVKEILKRNPVNLTDK DGNTALMIASKEGHIHIVEELDDACTYVNIIPDR SG  
 271 DTVLIGAVRGHVEIVRAILQKYADIDIRGQ DNKTALYWAVEKGNATMVRDILQCNPDTETCTK DGETPLIKATKMRNTEVVELLLDKGA  
 361 KVSADV KGDTPHVAIRGRSRRLAELLRLNPKDGRLLYRPN KAGETPNYNDCHSHQKSILTOIFGARHLSPTETDGMGLGYDLYSSALAD  
 451 ILSEPTMQPPICVGLYAAQWSSKSFLLKKLEDEMKTTFAGQQTETPLFQFSMLIVFELTLLCGGLGVFAFTVDTNLHIAISLSFLALIYIF  
 541 FIVIVYQGRREGESWNWAWALSRLARHIGYLELLFKLMFVNPPELAEQTTKALPVRELETTDXNRLSSVGGGETSLAEMIATLSDACEREF  
 631 GLATRLFRVLRTESQGGKKWKRTCCPLPSFVIFLFI VGCIIAGITLLAIFRVDPRKHLTVNAILISIASVUGLAFVLNCRTWQVLDLSL  
 721 NSQKRRLHSAASKLHKLKXSEGFNKVLKCEVELMARMAKTIDSTQNTQLRVIIDGLDACEQDKVLQMLDTRVLFKSGPTIAIFASDPH  
 811 IIAKAINQNLNSVLRDSNINGHDYMRNIVHLPVFLNSRGLSNARKFLVTSATNGDITCSDTGTQEDTDRRVSONSLGEMTKLGSKTALN  
 901 RRDYRRRQMQRTITRQM SFDLTLLVTEDFWSDISPTMRRLLNIVSVTGRLLRANQITFNWDRLASWINLTEQWPYRTSWLILYLEET  
 991 EGLPDQMTLKTIERISKNIPTTKDVEPLLEIDGDIRNTEVFLSSRTPVLVARDVKTFPLPCTVNLDPKLRETIADVRAAREQINIGGLAY  
 1081 PPLPLHEGPPRPSPGYSPASVCSASFNPGVSPQPHSSYSGLSGPQHPFYNR PFFAPYLYTPRYYPGGSQHLISRSSVKASLP  
 1171 RDQNNGLPCDSGFNKQRQSRPCHRLQTVTFNDRGRCCKLRQIEGLDQSMMPQYCTTIKKANINGRVLSQCNIDELKKEMAMNFGDWH  
 1261 LFRSMVLEMRSVESQVVPEDPRFLNENSSAPVPHGESARRPSHTPLTELSSQTPYTLNFSFEELNTLGLDEGAPRHSNLSWQSQTRRT  
 1351 PSLSSLSNQDSSIEISKLTDKVQAEYRDAYREYIAQMSQLEGGTSGSTISGRSSPHSTYIYQSSSGGSIHSTIEQERKEGELKQEDGR  
 1441 KSLMKRGDVIDYSSSGVSTNEASPLDPTTEDEKSDQSGSKLLPGKKSSERPFLQTDLLKGGGLRYQKLPDSEDESGTEESDNTPLL  
 1531 KDDKDKAEGKAERVCKSPHSAEPIRTFIKAKKEYLSDALLDKDSSDSGVRNENSSPNHSLHNEAADDSQLKANKLIELEDEGHSGKRG  
 1621 MPHSLGLODPTIARMSCSEDKSPSECSLIASSPEESWPACQKAYNLNRTPTSTVTLNNNTAPTNRANQNFDEIEGIRETSQVILRPGP  
 1711 SPNPTAVQENENLKSMAHKRSQRSSSYTRLKSDASELHAASSESTGFGEEESIL

Figure 2 (page 1 of 2)

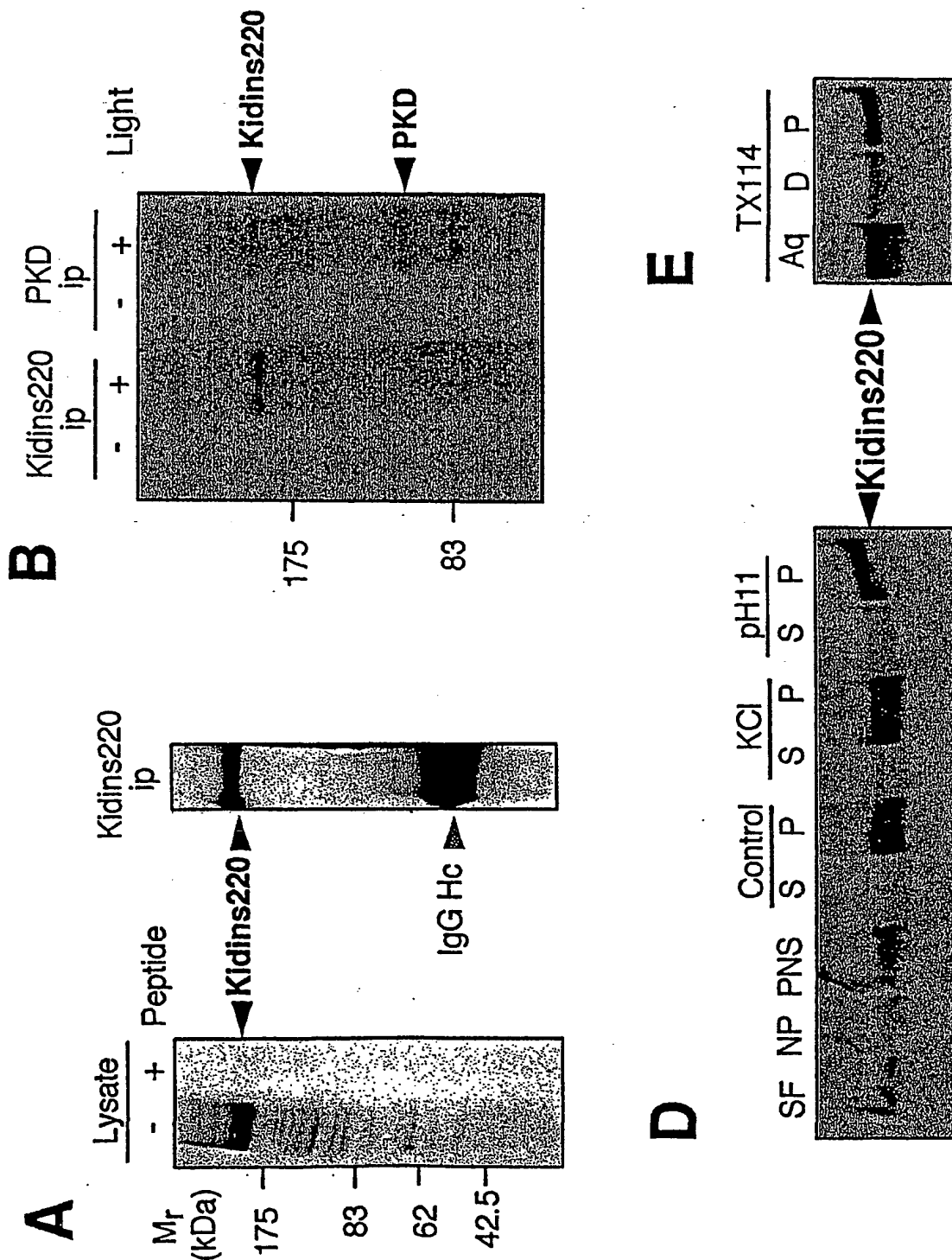


Figure 2 (page 2 of 2)

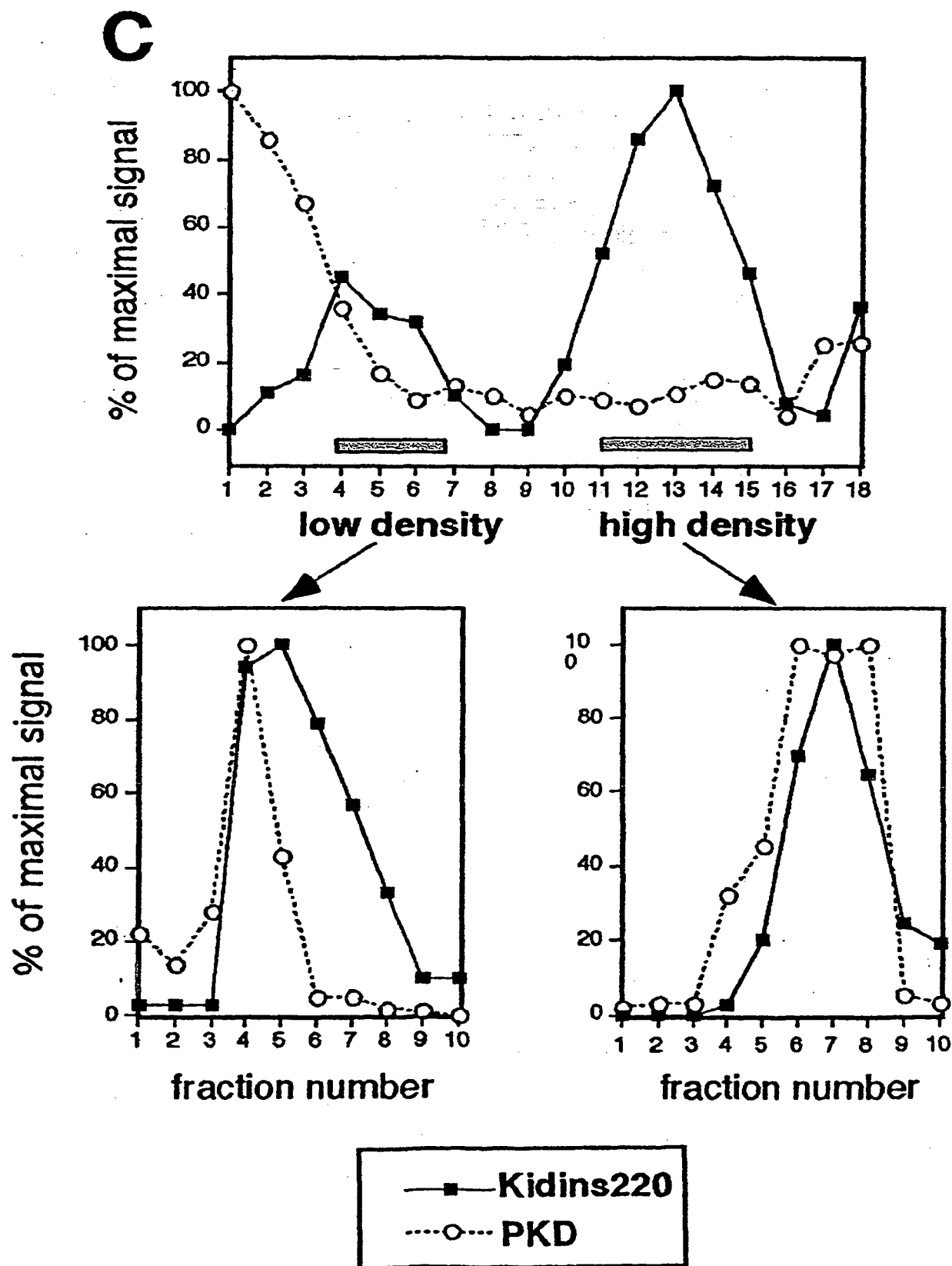


Figure 3

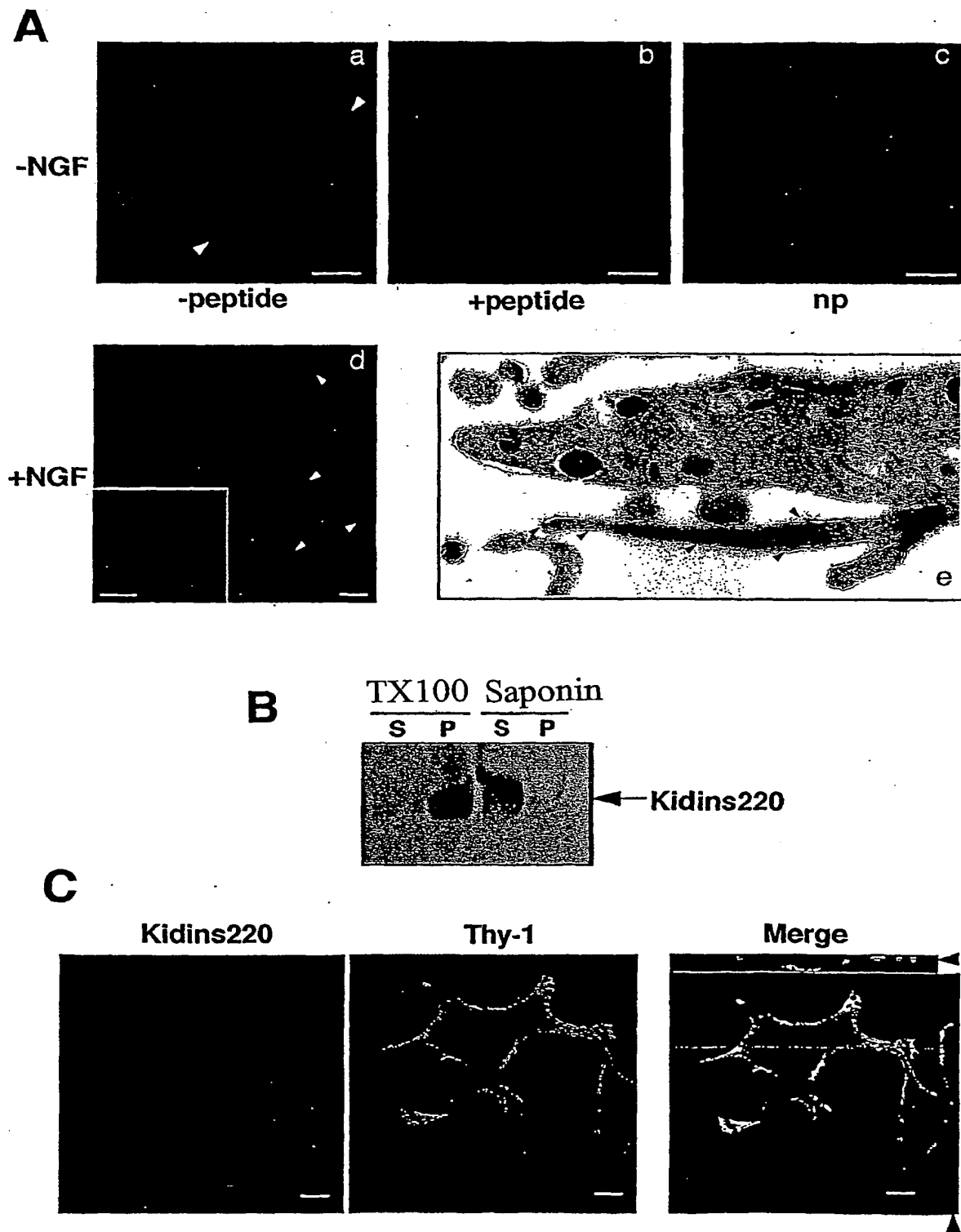


Figure 4

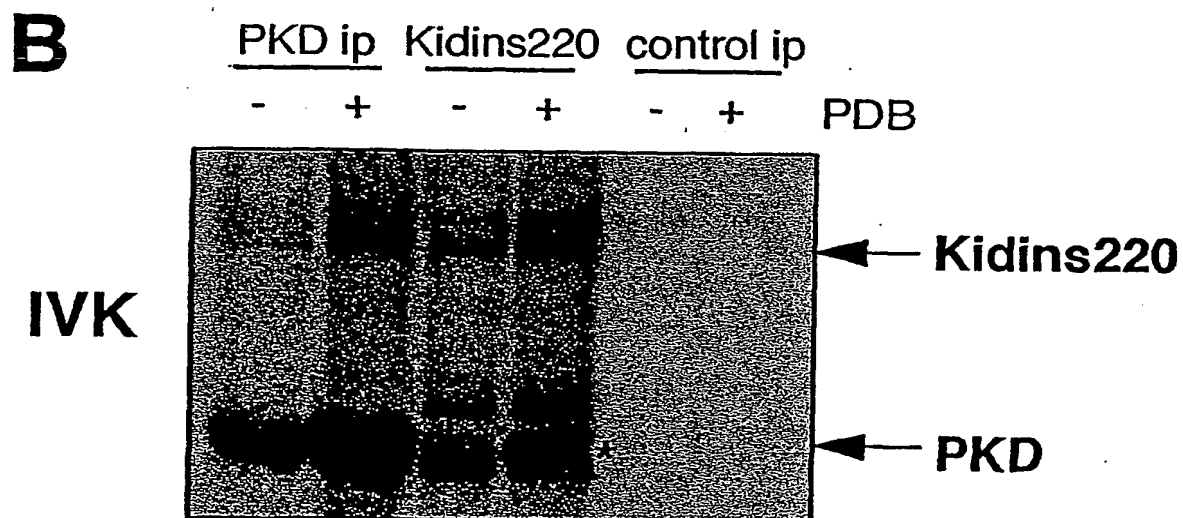
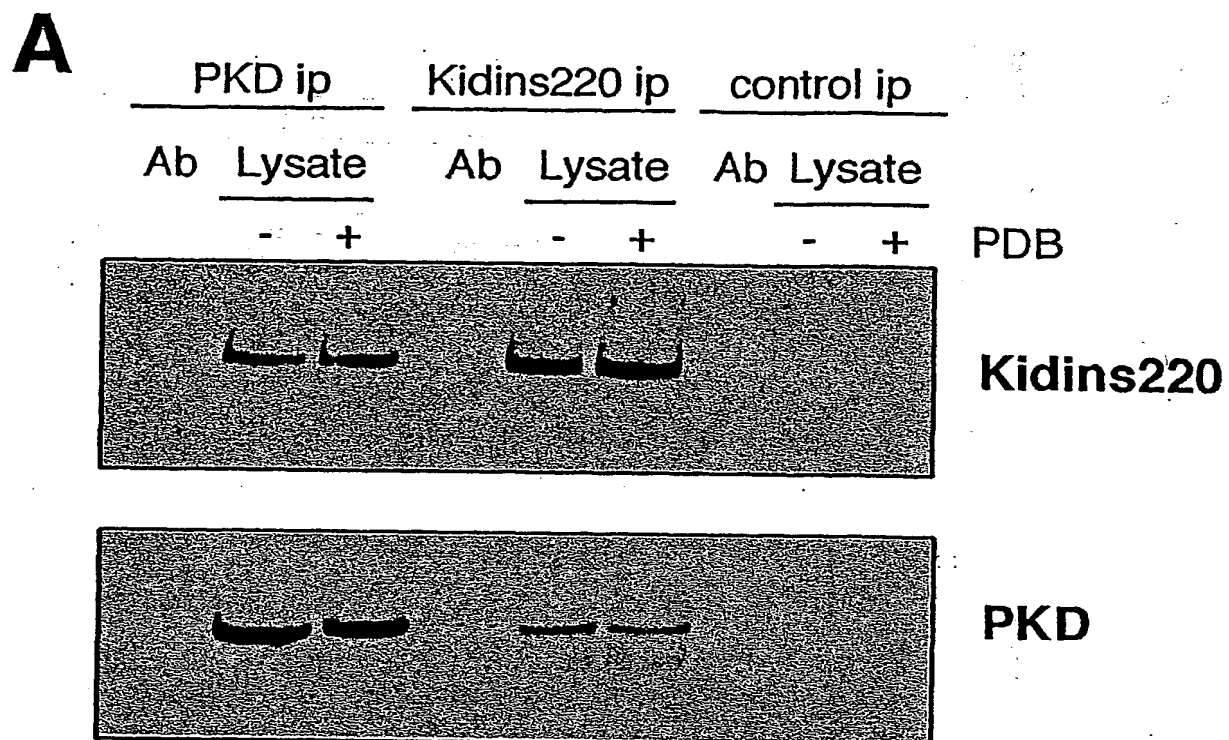
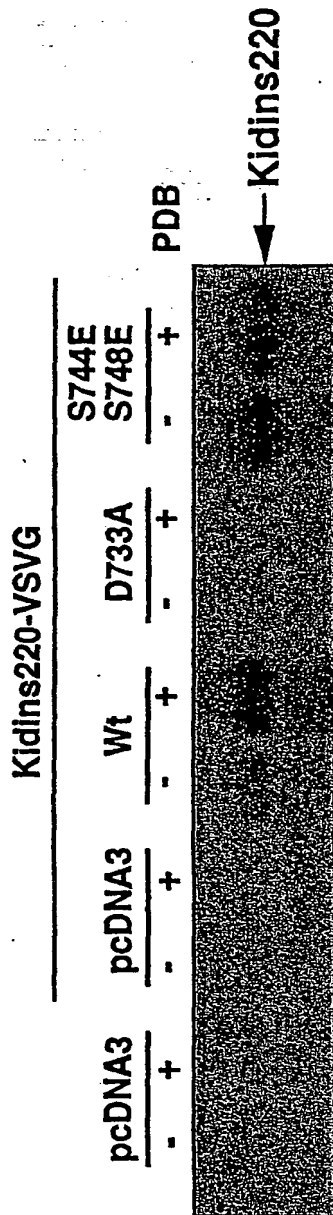


Figure 5 (page 1 of 2)

**A**



**B**

Kidins220-Ser919 908-RQMQRITRQM**S**FDLTK-924  
 Kidins220-Ala919 908-RQMQRITRQM**A**FDLTK-924  
 Kidins220 C-Term 1747-CAASSESTGFGEERSIL-1763



Figure 5 (page 2 of 2)

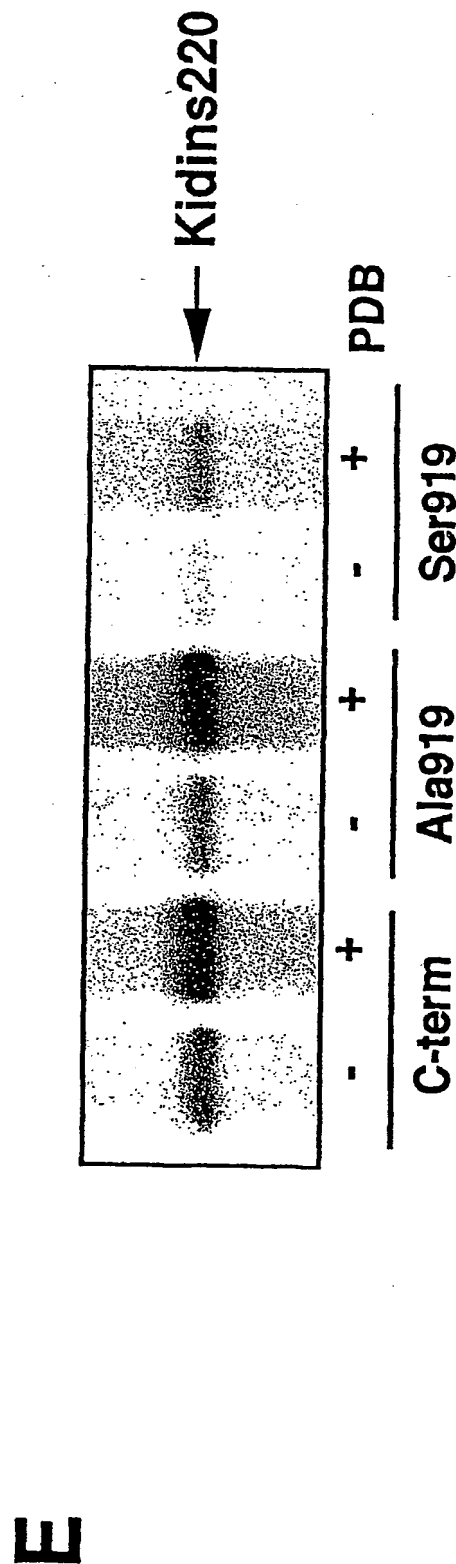
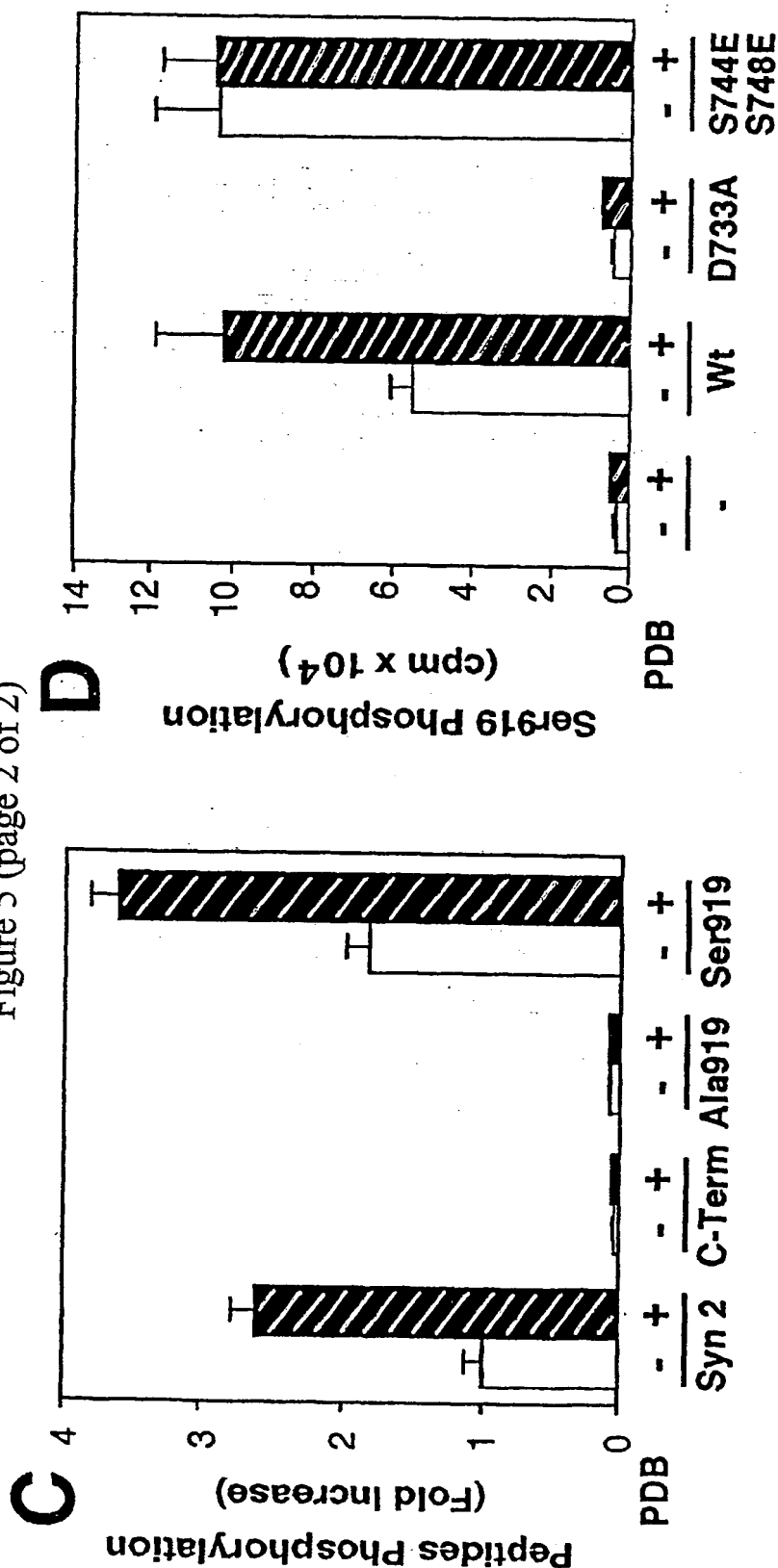
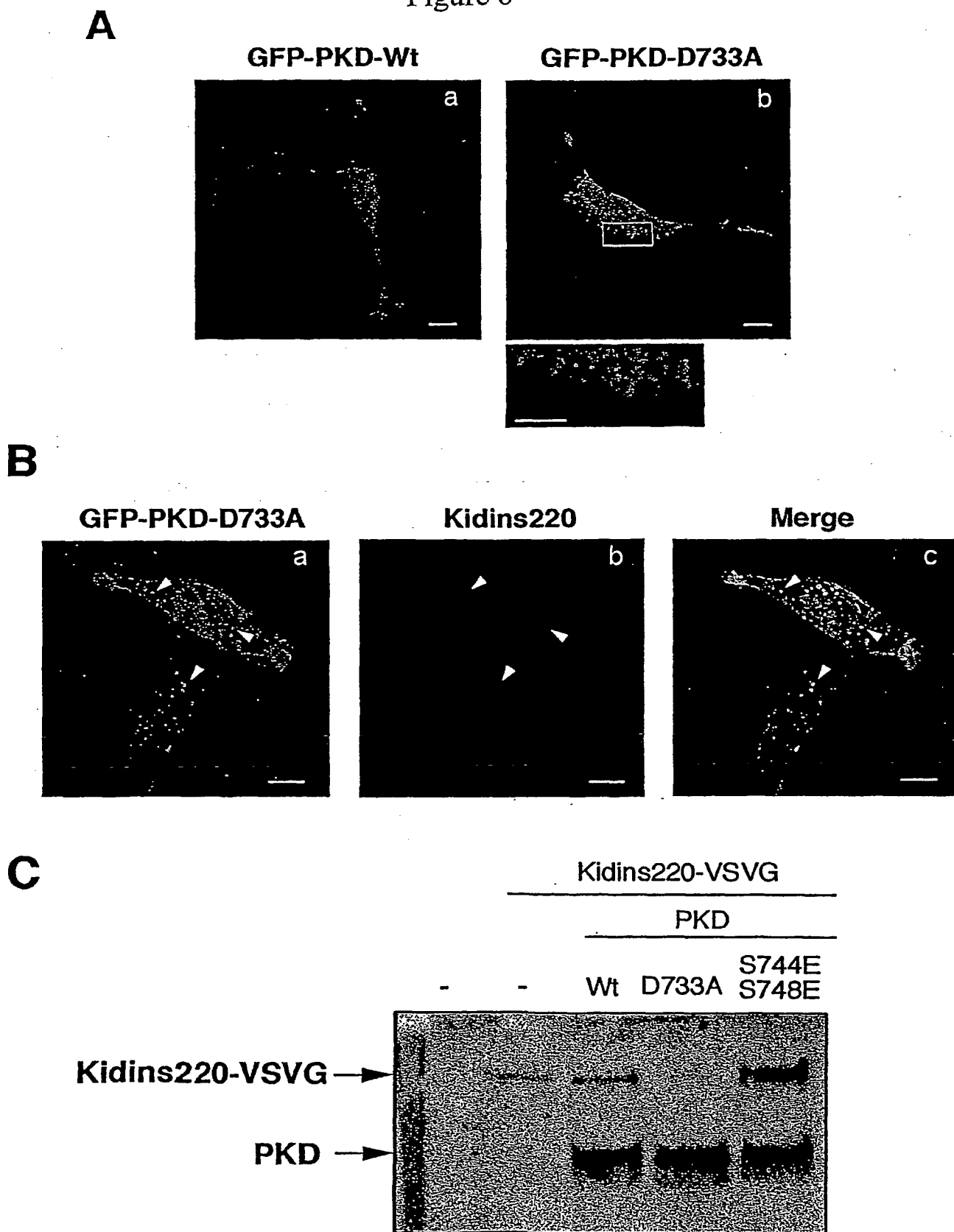


Figure 6



## Figure 7 (page 1 of 3)

## 5' untranslated

gtcctctgacaaggtctgcgccctgtgagctggctctacagagccgagggcgaggtc  
ctcagccaggccgagatacgggtgaattaaagtcggcaatagatttctgtgaaa

## ORF

ATGTCAGTTCTTATATCACAGAGTGTGATAAATTATGTGGAAGAAGAGAATATTCCTG  
CTCTGAAAGCTCTTCTTGAGAAGTGCAAAGATGTAGATGAGAGGAACGAGTGTGGCCA  
GACGCCGCTGATGCTGGCTGCGGAGCAGGGCAATGTGGAATAGTGAAAGAACTGTTA  
AAGAATGGAGCCAACTGCAATCTGGAAGATCTGGATAACTGGACTGCCCTTATATCGG  
CATCTAAGGAGGGGCACATCCACATTGTGGAGGAGCTGCTGAAGAGCGGGGCCAGTCT  
GGAGCACCGAGACATGGGTGGATGGACAGCTCTCATGTGGGCGTGCTACAAAGGCAGG  
ACCGACGTGGTTCGAGCTGCTTCTTCTCATGGTGCCAACCCAAGTGTCACTGGTTTGC  
AGTATAGTGTCTACCCAATCATCTGGGCAGCGGGCAGAGGCCATGCCGACATAGTGCA  
TCTTCTACTGCAGAATGGTGCGAAAGTCAATTGCTCCGACAAGTACGGAACAACCCCC  
TTGGTATGGGCTGCCCGAAAGGGTCATTTGGAATGTGTGAAGCACTTGTGGCCATGG  
GAGCTGATGTTGATCAAGAAGGAGCCAATTCGACGACTGCACTTATCGTGGCAGTGAA  
AGGAGGCTACACACAGTCCGTGAAAGAGATTTTGAAAAGGAATCCAAATGTGAACCTA  
ACTGACAAGGATGGAAACACAGCTCTGATGATCGCATCGAAGGAGGGACATATTGAGA  
TTGTGCAGGACCTGCTGGACGCCGGAACATACGTGAACATACCTGACAGGAGTGGGGA  
TACTGTATTGATCGGCGCTGTGAGAGGTGGTCATGTTGAAATTGTTTCGAGCACTTCTC  
CAAAAATATGCTGATATAGACATTAGAGGGCAGGACAATAAGACCGCGCTGTATTGGG  
CCGTTGAAAAGGGAAATGCGACAATGGTGAGGGGATATCCTGCAGTGCAATCCGGACAC  
GGAAATATGCACGAAGGATGGTGAGACACCACTGATAAAAGCCACCAAGATGAGAAAT  
ACTGAAGTGGTGGAGCTGCTGCTGGATAAGGGGGCTAAGGTGTCTGCTGTGGACAAGA  
AAGGGGACACTCCCCTGCACGTTGCCATCCGAGGGAGGAGCCGGAGACTGGCCGAACT  
TCTTTTACGGAACCCCAAAGATGGACGGTTACTCTACAGACCCAACAAAGCAGGCGAG  
ACTCCCTACAACATTGACTGCAGCCACCAGAAAAGCATTTTAACTCAAATATTTGGAG  
CCAGACACTTGTCTCCTACGGAAACAGATGGCGACATGCTTGGTTATGATTTGTATAG  
CAGTGCCCTGGCCGATATTCTCAGTGAACCCACCATGCAACCACCCATTTGTGTGGGC  
TTGTATGCACAGTGGGGAAGTGGGAAGTCTTTCTTACTCAAGAACTAGAAGATGAGA  
TGAAGACGTTTGCAGGACAGCAGACCGAGCCCCTCTCCAGTTCTCGTGGCTCATAGT  
GTTCCCTCACCTGCTGCTGTGTGGCGGGCTCGGCCCTAGTGTGTTGCTTTCACGGTTGAT  
ACAAACCTCGCCATAGCCATTTCAATTGAGTTTCTTGGCTCTCATATATATATTCTTCA  
TCGTCAATTTACTTCGGTGGACGGCGGGAAGGAGAGAGTTGGAACCTGGGCCTGGGCTCT  
CAGTACCAGACTGGCAAGACACATTGGGTACCTGGAGCTCCTCTTCAAACCTGATGTTT  
GTGAACCCACCTGAACTTGCGGAGCAAACCTACAAAGGCTTTACCTGTGAGGTTTTTGT  
TTACAGATTACAATAGACTCTCCAGTGTGGGAGGAGAACTTCCTTGGCTGAAATGAT  
CGCGACCTCTCAGATGCTTGTGAGAGAGAATTCGGCTTTTTTGGCTACCCGTCTTTTTT  
CGAGTCCTCAGGACTGAAGAGTCTCAGGGTAAAAAGAAATGGAAAAGAACATGCTGCC  
TCCCCCTCTTTTGTTCATCTTTCTTTTATTGTTGGCTGCATTATTGCTGGGATCACCCCT  
TCTGGCTATATTCAGAGTTGACCCGAAACATCTGACAGTGAATGCTATCCTCATTCT  
ATCGCATCTGTCGTGGGGTTGGCCTTCGTGCTCAATTGTGCGAACATGGTGGCAAGTGC  
TGGACTCTCTCCTGAATTCTCAGAGAAAACGCCTCCATAGTGCTGCCTCCAAATTACA  
CAAGTTAAAAAGTGAAGGATTCATGAAAGTCCTGAAGTGTGAGGTGGAGCTGATGGCC  
AGGATGGCAAAAACCATTTGACAGCTTCACTCAAACACAGACAAGGCTGGTGGTTATCA

10/79

SUBSTITUTE SHEET (RULE 26)

## Figure 7 (page 2 of 3)

TTGACGGCCTGGATGCCTGCGAGCAGGATAAAGTCCTTCAGATGCTGGACACTGTCCG  
AGTTCTGTTTTCAAAGGGCCCTTTTATTGCCATTTTTTGCAAGTGATCCGCATATTATC  
ATAAAGGCCATCAACCAGAACCTGAATAGTGTGCTTCGTGACTCAAATATAAACGGAC  
ATGACTACATGCGCAATATAGTTCACTTACCAGTTTTTCTTAATAGTCGTGGGCTTAG  
CAATGCAAGGAAATTCCTTGTAACCTCAGCAACAAATGGGGACATTACATGCTCAGAC  
ACCACAGGAACACAGGAGGACACTGACAGAAGAGTTTCACAGAACAGCCTTGGGGAGA  
TGACAAAGCTTGGGAGCAAAACGGCTCTCAACAGAAGGGGACACTTACCGCAGAAGACA  
GATGCAGAGAACCATCACGAGGCAGATGTCCTTTGATCTCACAAAGCTGCTGGTCACT  
GAGGATTGGTTCAGTGACATCAGCCCTCAGACCATGAGACGGCTACTCAATATTGTTT  
CTGTGACAGGGCGGCTACTAAGAGCCAATCAGATTACTTTCAACTGGGACAGGCTAGC  
CAGCTGGATCAACCTCACCGAGCAGTGGCCGTACCGGACGTCTTGGCTTATACTATAT  
CTGGAAGAGACTGAAGGTCTCCCTGATCAGATGACCTTAAAGACCATCTATGAGAGAA  
TATCGAAGAATATCCCAACAACCTAAAGATGTTGAGCCGCTTCTTGAAATTGATGGAGA  
TATAAGAACTTTGAAGTCTTCTTGTCTTCAAGGACTCCAGTTCTCGTGGCTCGAGAT  
GTAAAGACCTTTTTTGCCGTGCACAGTAAACCTGGATCCCAAACCTTCGGGAGACCATCG  
CAGATGTCCGAGCTGCAAGAGAGCAGATCAACATAGGAGGCCTGGCCTATCCCCCGCT  
CCCTCTGCACGAAGGCCCTCCTCGGCCACCTTCTGGGTACAGTCAGCCTGCATCTGTC  
TGCTCTTCAGCATCTTTCAACGGGGCCGTTCCAGGCGGGGTTGTGTCTCCACAGCCCC  
ACAGCAGTTACTACAGCGGCTTGTCCGGACCCACGACCCCTTCTACAACAGGCCATT  
CTTTGCCCCATACCTTTACACGCCAAGGTATTACCCTGGCGGTTCCCAACATCTCATC  
TCACGTTTCATCAGTAAAAGCGAGTTTGCCCAGAGATCAGAACAATGGCCTACCGTGTG  
ACTCTGGGTTTTAACAAACAGAGACAGGGCAGCCGTCCCTGCCACAGGCAGCTCACTGT  
TACTGAGTTCAATGACCGTGGATGTCGTTGTGAGAACTGAGACAGATAGAAGGGCTG  
GACCAGAGCATGATGCCTCAGTACTGCACAACAATCAAAAAGGCGAACATAAATGGCC  
GGGTATTGTCTCAGTGTAACATTGATGAGCTGAAGAAAGAGATGGCTATGAACTTTGG  
AGACTGGCATCTCTTCAGAAGCATGGTCTTAGAAATGAGAAGTGTGGAAAGCCAGGTG  
GTCCCTGAAGACCCTCGTTTTCTCAATGAAAACCTCCAGTGCCCCAGTTTCTCATGGAG  
AGTCTGCTCGCCGCCCTTCTCACACTGAGCTGCCCCCTCACGGAGCTTTCCAGCCAGAC  
TCCATACACCCTGAACTTCAGCTTTGAAGAACTGAACACGCTTGGCCTTGATGAGGGT  
GCCCCCTCGTCACAGTAATCTAAGTTGGCAGTCACAACTCGCAGAACCCCAAGTCTCT  
CGAGTCTCAATTCCCAGGACTCCAGTATTGAAATTTCAAAGCTTACTGATAAGGTGCA  
GGCCGAGTATAGAGACGCCTATAGAGAGTACATTGCTCAGATGTCCAGTTAGAAGGG  
GGCACAGGGTCGTCCACAATAAGCGGCAGATCTTCTCCACATAGCACATACTACATAG  
GCCAGAGCTCTTCAGGGGGCTCCATCCATTCTACTCTAGAACAGGAAAGGGGGAAGGA  
GGGTGAGCTAAAGCAGGAGGATGGGCGGAAGTCATTTTTTAATGAAGAGGGGGGATGTC  
ATAGATTACTCCTCATCAGGGGTGTCCACTAATGAGGCCTCACCCCTGGACCCCATTA  
CTGAAGAAGATGAGAAATCCGACCAGTCAGGTAGTAAGCTACTCCCAGGCAAGAAATC  
CTCAGAACGGCCCAGTCTCTTCCAGACAGACTTGAAGCTTAAGGGGGGTGGTCTTCGC  
TACCAGAAACTGCCCAGTGATGAAGATGAGTCGGGGACAGAAGAGTCAGATAACACCC  
CACTGCTCAAAGATGATAAGGACAAAAAGGCTGAAGGCCAAAGCAGAGAGAGTGTGCAA  
GTCCCCAGAGCACAGTGCCGAGCCCATCAGAACGTTTATTAAAGCAAAAGAGTATTTG  
TCCGATGCCCTCCTTGACAAAAAGGATTCTCAGATTCGGGAGTGAGGTCCAATGAGA  
GTTCTCCCAACCACTCTCTCCACAATGAGGCAGCAGACGACTCCCAACTGGAAAAGGC  
AAACCTCATAGAGCTTGAAGATGAGGGCCACAGTGGGAAGCGGGGGATGCCACACAGT  
CTGAGTGGCTTGCAAGATCCAATTATAGCTCGGATGTCCATTTGCTCGGAAGACAAGA  
AAAGCCCTTCCGAGTGCAGCCTGATTGCTAGCAGCCCTGAAGAAAGCTGGCCTGCGTG  
CCAGAAGGCCTACAATCTAAACCGAACGCCAGCACTGTGACTCTGAACAACAACACT  
GCACCCACCAACAGAGCCAATCAGAACTTCGATGAGATAGAGGGAATCAGGGAGACGT

11/79

SUBSTITUTE SHEET (RULE 26)

## Figure 7 (page 3 of 3)

CCCAGGTCATTCTGAGGCCTGGCCCCAGTCCCAACCCCAACTGCTGTGCAGAATGAAAA  
CCTGAAGAGCATGGCGCACAAAGCGAAGCCAGCGCTCCAGTTACACAAGGCTCTCGAAA  
GATGCATCTGAGCTACATGCAGCCTCCTCGGAGAGCACAGGCTTTGGAGAAGAGAGAG  
AGAGCATTCTC

**Stop codon and 3' untranslated**

tgaggaaggagaagtgaaaagagtgtgacagactttgagttgaccattcctcattactcatca  
tgtttatgcatcacagacacctgataggtggtcaggcactggaagcagacaggaaagcctgaa  
agcacagtcaccagtcattttcattgtcaccgtcatcccaggcactttaaagtctcactgcca  
ctattacacctttcctggcacacagctaggcagagcgatgagagcaggctgaagacaaaatag  
gtcaaaagtgtattttaagttgaaagggcttaaatcacaaggaaaaaaacttagagatctcct  
ttataaactattcatgtaccatttgagggaagatacagtggttagatgaagcagaaaccaattt  
tattgttagtggttaattcttggtgcagatttataagtttttagagttagcccagagactaaaagt  
gaatacttagcaaatggatagccagtgttctatataggaatcattgcttttcagaggggcttaa  
agtttaagttagaaaaatataactcaagaaggcgataaaaagctgatgagaaagttagtagcag  
aacccaaagccgtgctgggcccgggtgactcattagcagaggaggagggggaggggcagtattt  
cctgggatactctctccagaccagcctggcttctgacatcatccacctgtgcctcctcaaaacc  
gtcttagtctgttctgcaactcttaagtacatacctaactcagctcatggctaaggaaaaaa  
aattaaagtgtcctgggtgattaaactctggacctcccacatctaagtcaggagttgacaaac  
tgcacccccagcccaaacatctgttcttcagagatctggaaagacatgggcatgaagagtcca  
gtaaactcttaatgaagggtgtcgggaagtgtttcccttaaagggggttagtgacagtacacta  
gacaagtgacagaatttagtttgctgacattaaatcactgccttctgggcccgaatagtttag  
tatctgttccctaaatcagcatttgagttctgcgggggttacaaaatggaggcatggttttaag  
aaggaatttgaaagcatccagggaagcaagaacagagtcggagctggtgtgaccatgacactt  
cccattattgttagcaaaagctaggagaatgaccacatacatttaagtatcatgtaggctagag  
cgtgcctgtcatcatcatgcattgagaccgtttgtcctttgattatagtggtttatgtagac  
cagtccttagcagtatgtcctgagccatgggcaggattgtagatggcgagggtggcagggttagt  
gctcgttgagatgtccatctctgactggacctgtgtgggttactttagctgctccgtgaaac  
actgccttgctgtttgagtgttaattgagcattttcatgaagtgtatatgttgcatgatcttta  
acaggtatttttagacattaagaacagcatgatattactgacatatgcattagtagtatgcag  
tatcatgtaatgggtacttcagaggcagtttggtttctttccagggtgagtagttatagctcac  
agtcatgggaatcctttattcacactgtctgtgaagtcgatgctggagtgccaaatgcacacg  
gagactaggaacaataaagacttgcatgccattctgactcatggagttttattcaatagagca  
atgataattcccattctgattaaaaattaaaacacggaa

12/79

SUBSTITUTE SHEET (RULE 26)

Figure 8 (page 1 of 4)

	90		10		20		30		40		50		60		70		80
--	----	--	----	--	----	--	----	--	----	--	----	--	----	--	----	--	----

Clone 1 (300bp/101aa DELETION)

aa1131 [1130]

GPQHFFYNR [-

----- ] ANINGRVLSQ 1250 [1249]

bp 3391 [3499]

GGACCCAGCAGCACCCCTTCTACAACAGG-----

-----GCGAACATAAATGCCCGGTATTGTCTCAG

Clone 2 (171bp/57aa DELETION)

aa1131 [1130] GPQHFFYNR [-

-----AAVP

ATGSSLLLS MTVDVVCEKL RQIEGLDQSM MPQYCTTIKK ] ANINGRVLSQ 1250 [1249]

Figure 8 (page 2 of 4)

bp 3391 [3499]  
 GGACCCAGCACCCCTTCTACAAACAGG-----  
 -----  
 --GCAGCCGTCCCTGCCACAGGCAGCTCACTGTTACTGTAGTTCAATGACCCGTGGATGTCGTATGTGAGAACTGAGACAGATAGAGGGCTGGACCCAG  
 AGCATGATGCCTCAGTACTGCACAAACAATCAAAAAGCGCAACATAAATGGCCGGGTATTGTCTCAG

clone 3 (57bp/19aa DELETION)  
 aa1131 [1130]  
 GPQHFFYNR [P FFAPYLYTPR YYPGGSQHLL SRSSVKASLP RDQNNGL--- -----AAVP ATGSSLLLS  
 MTVDVVCEKL RQIEGLDQSM MPQYCTTIKK ]ANINGRVLSQ 1250 [1249]

bp 3391 [3499]  
 GGACCCAGCACCCCTTCTACAAAGGCCATCTTTGGCCCCATACCTTTACAGCCCAAGGTATTACCTGGCGGTTCCCAACATCTCATCTCAGTTC  
 ATCAGTAAAGCGAGTTTGCCCGAGAGATCAGAACAAATGGCCTA-----  
 --CAGCCGTCCCTGCCACAGGCAGCTCACTGTTACTGTAGTTCAATGGACCCGTGGATGTCGTATGTGAGAACTGAGACAGATAGAGGGCTGGACCCAG  
 AGCATGATGCCTCAGTACTGCACAAACAATCAAAAAGCGCAACATAAATGGCCGGGTATTGTCTCAG

14/79

SUBSTITUTE SHEET (RULE 26)

Figure 8 (page 3 of 4)

Clone 4 (9bp/3aa DELETION + 16aa VARIATION)

aa1131 [1130]

GPQHPFYNR [P FFAPYLYTPR YYPGGSQHLL SRSSVKASLP RDQNNGLPCD SGFNKQRQVQ KIQ---AAVP ATGSSLLLS

MTVDVVCEKL RQIEGLDQSM MPQYCTTIKK ] ANINGRVLSQ 1250 [1249]

bp 3391 [3499]

GGACCCAGCACCCCTTCTACAAAGGCCATTCTTTGCCCATACCTTTACAGCCCAAGGTATTACCTGGCGGTTCCCAACATCTCATCTCACGTTTC  
ATCAGTAAAGCGAGTTTGCCAGAGATCAGAACAAATGGCCCTACCGTGTGACTCTGGGTTTAAACAAACAGACAGGTACAGAAACTCCAG-----  
--GCAGCCGTCCCTGCCACAGGCAGCTCACTGTTACTGAGTTCAATGACCCGTGGATGTCGTGAGAAACTGAGACAGATAGAAAGGCTGGACCAG  
AGCATGATGCCTCAGTACTGCACAAACAATCAAAAAGGCGAACATAAATGGCCGGGTATTGTCTCAG

Clone 5/1 (24bp/8aa DELETION + 31aa VARIATION)

aa1131 [1130]

GPQHPFYNR [P FFAPYLYTPR YYPGGSQHLL SRSSVKASLP RDQNNGLPCD SGFNKQRQ---GSRP CHRQLTVTEF

NDRGCRCEKL RQIEGLDQSM MPQYCTTIKK ] ANINGRVLSQ 1250 [1249]



Figure 8 (page 4 of 4)

bp 3391 [3499]

GGACCCAGCAGACCCCTTCTACAAACAGGCCATTCTTTGGCCCATACCTTTACAGCCCAAGGTATTACCTGGCGGTTCCCAACATCTCATCTCAGTTTC  
 ATCAGTAAAGCGAGTTTGGCCAGAGATCAGAAACAATGGCCCTACCGTGTGACTCTGGGTTTAAACAAACAGAGACAG-----  
 --GGCAGCCGTCCTGCCACAGGCAGCTCACTGTTACTGAGTTCAATGACCGTGGATGTCGTTGTGAGAAACTGAGACAGATAGAAGGGCTGGACC  
 AGAGCATGATGCCTCAGTACTGCACAAC AATCAAAAAGCGCAACATAAATGGCCGGTATTGTCTCAG

Clone 5/2 (24bp/8aa DELETION + 11aa VARIATION)

aa1131 [1130]

GPQHPFYNR [P FFAPLYTPR YYPGGSQHLLI SRSSVKTSLP RDQNNGLPCD SGFNKQRQ-- -----AAVP ATGSSLLLS

MTVDVVCEKL RQIEGLDQSM MPQYCTTIKK ] ANINGRVLSQ

Clone 6 (19aa VARIATION)

aa1131 [1130]

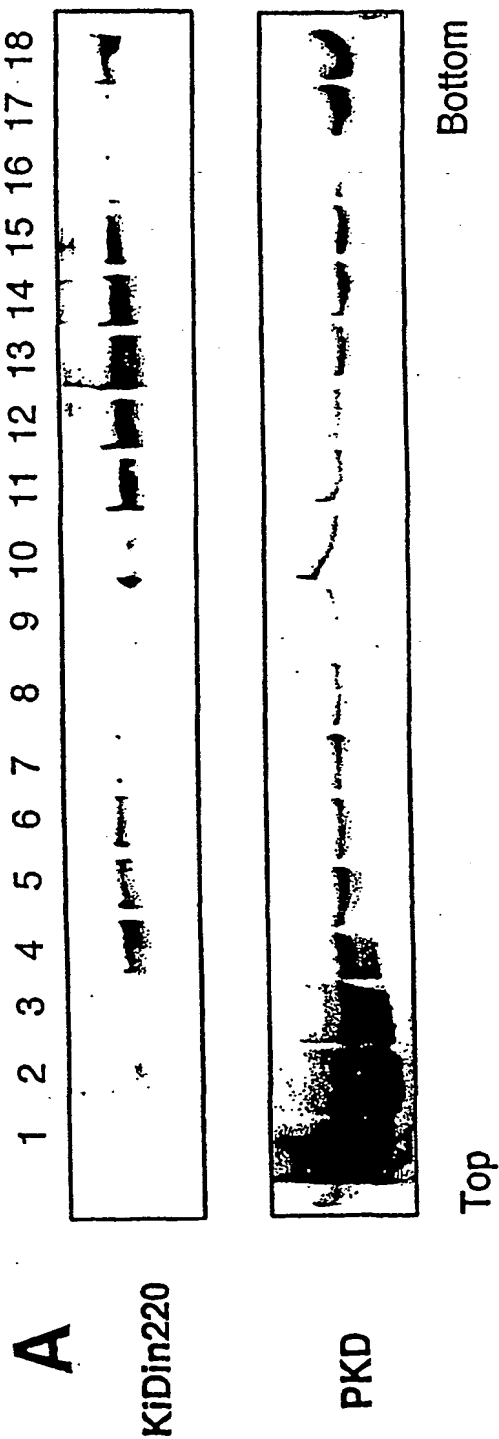
GPQHPFYNR [P FFAPLYTPR YYPGGSQHLLI SRSSVKASLP RDQNNGLEVI TEDAAEGLPS PTDSSRAAVP ATGSSLLLS

MTVDVVCEKL RQIEGLDQSM MPQYCTTIKK ] ANINGRVLSQ 1250 [1249]

bp 3393 [3499]

GGACCCAGCAGACCCCTTCTACAAACAGGCCATTCTTTGGCCCATACCTTTACAGCCCAAGGTATTACCTGGCGGTTCCCAACATCTCATCTCAGTTTC  
 ATCAGTAAAGCGAGTTTGGCCAGAGATCAGAAACAATGGCCCTAGAAGTAATCAAGAGATGCCGTGAGGGCTTCCCTCACCACAGACTCCTCGC  
 GGCAGCCGTCCTGCCACAGGCAGCTCACTGTTACTGAGTTCAATGACCGTGGATGTCGTTGTGAGAAACTGAGACAGATAGAAGGGCTGGACCAG  
 AGCATGATGCCTCAGTACTGCACAACAAATCAAAAAGCGCAACATAAATGGCCGGTATTGTCTCAG

Figure 9



**B**



Figure 10 (page 1 of 2)

```

lalign output for AL133620 vs. AB033076
AL133620      1031 aa vs.
AB033076      1216 aa
scoring matrix: BLOSUM50, gap penalties: -70/-4
29.6% identity;      Global alignment score: -1573
-----
AL133620
AB033076
MSVLISQSVINYVEENIPALKALEKCKVDVERNECGQTPLMIAAEOGNLEIVKELIKNGANCNLEDLDNWTALISASKEGHVHIVEELLKCGVNLHHRDMGGWTALMWACYKGRTDVV 120
-----
ELLLSHGANSVVTGLQYSVYPIIWAAGRGHADI VHLLQLQNGAKVNCSDKYGTTPLVWAARKGHLECVKHLAMGADVDOEGANSMTALIVAVKGGYTQSVKEILKRNPNVNLTDKDGNTA 240
-----
LMTASKEGHEIVQDLLDAGTVYNIPDRSGDTVLIGAVRGHVEIVRALLQKYADIDIRGQDNKTALYWAVEKGNATWVRDILQCNPDPTEICTKQGETPLIKATKMRNIEWELLDDKGA 360
-----
KVSAVDKKGDTPLHIAIRGRSRKLAELLRNPKDGRLLYRPNKAGETPNIDCSHQSIILTIQIFGARHLSPTETDGM LGDYSSALADILSEPTMQPPICVGLYAQWGSKSFLLKKL 480
-----
EDEMKTFAQQQIEPLFQFSWLVFLTLTLGGLGLLFAFTVHPNLGIAVSLSFLALLYIFFTVIYFGRRREGESWNWAVLSTRLARHIGYLELLKLMFVNPPELPEQTTKALPVRFLF 600
-----
WLVFLTLTLGGLGLLFAFTVHPNLGIAVSLSFLALLYIFFTVIYFGRRREGESWNWAVLSTRLARHIGYLELLKLMFVNPPELPEQTTKALPVRFLF 100
-----
TDYNRLSSVGGETSLAEMIATLSDACEREFGLATRLFRVFKTEDTQKKKWKKTCCLPSSFVIFLFIIGCIISGITLLAIFRVPDPKHLTVNAVLIASVVGGLAFVLCNRTWQVLDSSL 720
-----
TDYNRLSSVGGETSLAEMIATLSDACEREFGLATRLFRVFKTEDTQKKKWKKTCCLPSSFVIFLFIIGCIISGITLLAIFRVPDPKHLTVNAVLIASVVGGLAFVLCNRTWQVLDSSL 220
-----
NSQKRRLHNAASKLHKLKSEGFMKVLKCEVELMARMAKTIIDSFTQNTQRLVVIDGLDACEQDKVLQMLDTRVLFSGKPFIAIFASDPHIIKAINQNLSVLRDSNNGHDYMRNIVH 840
-----
NSQKRRLHNAASKLHKLKSEGFMKVLKCEVELMARMAKTIIDSFTQNTQRLVVIDGLDACEQDKVLQMLDTRVLFSGKPFIAIFASDPHIIKAINQNLSVLRDSNNGHDYMRNIVH 340

```

[illegible]

## Figure 11 (page 1 of 2)

## Human Kidins Protein Sequence

Length: 1715 June 29, 2000 14:41 Type: P

1 MSVLISQSVI NYVEEENIPA LKALLEKCKD VDERNECGQT PLMIAAEQGN  
51 LEIVKELIKN GANCNLEDLD NWTALISASK EGHVHIVEEL LKCGVNLEHR  
101 DMGGWTALMW ACYKGRTDVV ELLLSHGANP SVTGLQYSVY PIIWAAGRGH  
151 ADIVHLLLQN GAKVNCSDKY GTTPLVWAAR KGHLECVKHL LAMGADVDQE  
201 GANSMTALIV AVKGGYTQSV KEILKRNPV NLTDKDGNTA LMIASKEGHT  
251 EIVQDLLDAG TYVNIPDRSG DTVLIGAVRG GHVEIVRALL QKYADIDIRG  
301 QDNKTALYWA VEKGNATMVR DILQCNPDTE ICTKDGETPL IKATKMRNIE  
351 VVELLLDKGA KVS AVDKKGD TPLHIAIRGR SRKLAELLR NPKDGRLLYR  
401 PNKAGETPYN IDC SHQKSIL TQIFGARHLS PTETDG DMLG YDLYSSALAD  
451 ILSEPTMQPP ICVGLYAQWG SGKSFLKKL EDEMKT FAGQ QIEPLFQFSW  
501 LIVFLTLLC GGLGLLFAFT VHPNLGIAVS LSFLALLYIF FIVIYFGGRR  
551 EGESWNWAWV LSTRLARHIG YLELLKL MF VNPPELPEQT TKALPVRFLF  
601 TDYNRLSSVG GETSLAEMIA TLSDACEREF GFLATRLFRV FKTEDTQGKK  
651 KWKKTCCLPS FVIFLFIIGC IISGITLLAI FRVDPKHLTV NAVLISIASV  
701 VGLAFVLNCR TWWQVLDLL NSQRKRLHNA ASKLHKLKSE GFMKVLKCEV  
751 ELMARMAKTI DSFTQNQTRL VVIIDGLDAC EQDKVLQMLD TVRVLFSGKP  
801 FIAIFASDPH I I KAINQNL NSVLRDSNIN GHDYMRNIVH LPVFLNSRGL  
851 SNARKFLVTS ATNGDVPCSD TTGIQEDADR RVSQNSLGEM TKLGSKTALN  
901 RRD TYRRRQM QRTITRQMSF DLTKLLVTE WFS DISPQTM RRLNIVSVT  
951 GRLLRANQIS FNWDRLASWI NLTEQWPYRT SWLILYLEET EGIPDQMTLK  
1001 TIYERISKNI PTTKDVEPLL EIDGDIRNFE VFLSSRTPVL VARDVKVFLP  
1051 CTVNLDPKLR EIIADVRAAR EQISIGGLAY PPLPLHEGPP RAPSGYSQPP

20/79

SUBSTITUTE SHEET (RULE 26)

## Figure 11 (page 2 of 2)

1101 SVCSSTSFNG PFAGGVVSPQ PHSSYYSGMT GPQHPPFYNRG SGPAPGPVVL  
1151 LNSLNVDVAVC EKLKQIEGLD QSMLPOYCTT IKKANINGRV LAQCNIDELK  
1201 KEMNMNFGDW HLFIRSTVLEM RNAESHVVPE DPRFLSESSS GPAPHGEPAR  
1251 RASHNELPHT ELSSQTPYTL NFSFEELNTL GLDEGAPRHS NLSWQSQTRR  
1301 TPSLSSLNSQ DSSIEISKLT DKVQAEYRDA YREYIAQMSQ LEGGPGSTTI  
1351 SGRSSPHSTY YMGQSSSGGS IHSNLEQEKG KDSEPKPDDG RKSFLMKRGD  
1401 VIDYSSSGVS TNDASPLDPI TEEDEKSDQS GSKLLPGKKS SERSSLFQTD  
1451 LKLKGSGLRY QKLPSDEDES GTEESDNTPL LKDDKDRKAE GKVERVPKSP  
1501 EHSAEPIRTF IKAKEYLSDA LLDKKDSSDS GVRSSSESPN HSLHNEVADD  
1551 SQLEKANLIE LEDDSHSGKR GIPHSLSGLQ DPIIARMSIC SEDKKSPSEC  
1601 SLIASSPEEN WPACQKAYNL NRTPSTVTLN NNSAPANRAN QNFDEMEGIR  
1651 ETSQVILRPS SSPNPTTIQN ENLKSMTHKR SQRSSYTRLS KDPPELHAAA  
1701 SSESTGFGEE RESIL

Figure 12 (page 1 of 6)

RnKidins	-----MSVLISQSVINYVEEENIPALKALLEK-CKDVDERNEC
HsKidins	-----MSVLISQSVINYVEEENIPALKALLEK-CKDVDERNEC
DmKidins	MFKSRLKTNPGGSENINRFGDSMGLGHRALLQYIDNNDISGLRAILDSRHLTIDDRDEN
CeKidins	-----MVQTLRRPWQEAAASSAFAVASALPVT-----MNST
	: : : : : *
RnKidins	GQTPMLAAEQGNVEIVKELLKNGANCNLEDLDNWTALISASKEGHIHIVEELLKSG-AS
HsKidins	GQTPMLAAEQGNLEIVKELLKNGANCNLEDLDNWTALISASKEGHVHIVEELLKCG-VN
DmKidins	ATTVLVWVAGRGLTAFVREFLARGADVQAEEDLDNWTALLCASRNGHLDVQQLLDHG-AE
CeKidins	QIAELFEQVEHGTTELRCALTAEIS--ALRNANGESLLTVAVRSGNTAVAKQLAQLDPPDA
	: *: : : : : * : : : : *
RnKidins	LEHRDMGGWTALMWACYKGRTDVVELLLSHGANPSVTGLQYSVYPIIWAAGRGHADIVHL
HsKidins	LEHRDMGGWTALMWACYKGRTDVVELLLSHGANPSVTGLQYSVYPIIWAAGRGHADIVHL
DmKidins	VEHRDMGGWTSLMWAAARGHTELVRLLLDKGADGNAHG-NYHLGALLWAAGRGYKDIVEL
CeKidins	IDETDNEGWSALLNAAHCGHVDIVRLLLDNGASVDQPD-LMGWSPLMWAVYKNHLDVVVDL
	: : * **: : * : : : : * : : : : *
RnKidins	LLQNGAKVNCSDKYGTTPLVWAARKGHLECVKHLLAMGADVDAQNSTTALIVAVKGGY
HsKidins	LLQNGAKVNCSDKYGTTPLVWAARKGHLECVKHLLAMGADVDAQNSTTALIVAVKGGY
DmKidins	LVQRGAKNVGDKYGTTLVWACRRGNVEIVDTLLKAGANVDTAGMYSWTPLLVAAAGGH
CeKidins	LVN--AKVNACDKFGSTALIWAARKGHLPPVQVQLLNSGAEVDVAVGMYSSSTALMLATRGNF
	: : : * **: : : : : * : : : : *
RnKidins	TQSVKEILKRNPVNLTDKDGN TALMIASKEGHI EIVQD LLDAGTYVNI PDRSGDTVLIG
HsKidins	TQSVKEILKRNPVNLTDKDGN TALMIASKEGHEI VQD LLDAGTYVNI PDRSGDTVLIG
DmKidins	TDCVSSILEKKPNVNALDKDGMTALCIASREGFQDIAASLIAAGAYINIQDRGADTPLIH
CeKidins	IQVVELLLTREPNNVADQNGLTALGMAARDGYADICESLINSGAFNQCDRFGNWILTS
	: * : * : : : : * : : : : * : : : : *
RnKidins	AVRGHVEIVRALLQKYADIDIRGQDNKTALYWAVEKGNATMVRDILQCNPDTEICTKDG
HsKidins	AVRGHVEIVRALLQKYADIDIRGQDNKTALYWAVEKGNATMVRDILQCNPDTEICTKDG
DmKidins	AVKAGHRTVVEALLKKHADVDIQGKDRKTAITYAVEKGHTPIVKLLLATNPDLLESATKDG
CeKidins	AVRSGNAAIVRMILDKFADINQDSEKRTPLHLAIDKSFNDIAYILLEKKPNLELKNKDG
	: : : * : * : : : : : : : : : : : : : : *

Figure 12 (page 2 of 6)

[illegible]



Figure 12 (page 3 of 6)

[illegible]

Figure 12 (page 4 of 6)

RnKiddins	RNFEVFLSSR-----TPVLVARDVKFTELPCTVNLDPKLRETIADVRAAREQINIGGLAYPPP
HsKiddins	RNFEVFLSSR-----TPVLVARDVKFTELPCTVNLDPKLREI IADVRAAREQISIGGLAYPPP
DmKiddins	RKLD AFLQLH-----KSDLLVADLRIFLPFTINLDPYLRKVLKEDQQTIEDE--GSLVIOQA
CeKiddins	ENFEGFLDSKGIPSAERLTVGHVKKFVPCCTSSLDPYLRKLIRE-----
	*** ** : * . . : : * : * * . * * * * : : :
RnKiddins	LPLHEGPPRRPPSGYSQPASVCSSASFNGPFPFGVGVSPQPHSSYYSGLSGPHQPFYNRPPFF
HsKiddins	LPLHEGPPRAPSGYSQPPSPVCSSTSFNGPFAAGVSPQPHSSYYSGMTGPQHPPFYNR----
DmKiddins	RPSVSNTMRQFPAPTIVPSPQAYPPYQMFQNEYPANELRSRNLSTSTEPVTPLINSP----
CeKiddins	R--SKG-----LVDIEAQVGSAGMAIPPVARLLFS-----
	.. : : : : *
RnKiddins	APLYTPRYYPGGSQHLSRSSSVKASLPRDQNGNGLPCDSGFNKQRGSRPCHRQLTIVTEF
HsKiddins	-----GSGPAPGPPVLLNS
DmKiddins	-----SDSFGDDILQTKLTD
CeKiddins	-----DDLWTMSIDTPLVE
	..
RnKiddins	NDRGCRCEKLRQIEGLDQSMMPQYCTTIKKANINGRVLSCQNIDELKEMAMNFGDWHLF
HsKiddins	LNVD AVCEKLQIEGLDQSMPLPQYCTTIKKANINGRVLACQNIDELKEMMNFGDWHLF
DmKiddins	LTVEGVISLDDRIEDMKP-ALPKLAPVLRENAINGRVLKHCMPDLKSVLGLSFGHWELF
CeKiddins	MKLD AVVNLRKIDIPSN-RLDSILDRFYQLNLCGLVLATCPLPELKDSMQLPLGDWTLI
	. : * : . . : : : * * * * : : * : : * : *
RnKiddins	RSMVLEMRSVESQVVPEDPRFLNENSSAPVPHGESARRPSHTELPTELSSQTPYTLNFS
HsKiddins	RSTVLEMRNAESHVVPEDPRFLSESSGPAPHGEPARRASHNELPHTELSSQTPYTLNFS
DmKiddins	RLIITTLR-----ECERLPRKQORQQQPGALEAPSNVPMIKDVTDALMQPPRESLS
CeKiddins	RLIILETLKVFG-----SSPPGLRVDKRKALTLRDEDEEEEEIEEAAALNSEKERAPLLG
	* : : : . . . . .
RnKiddins	FEELNTLGLDEGAPRHSNLSWQSQTRRTTPSLSSLSNSQDSSIEISKLTDKVQAEYRDAYRE
HsKiddins	FEELNTLGLDEGAPRHSNLSWQSQTRRTTPSLSSLSNSQDSSIEISKLTDKVQAEYRDAYRE
DmKiddins	--RK-----N-----SV--S--HMEKQVTL EEQMICGTLQTLNEEAYED
CeKiddins	SVRAE-----QRRRSTIVQNA TELSIDHKCLMEKLSGMDLTETEGDVNEMHFSHFSS
	. : : : : . : : : . : : .

Figure 12 (page 5 of 6)

RnKidins	YIAQMSQLEGGTGSSTISGRSSPHSTYYIGSSSSGGSIHSTLEQERGKEGELKQEDGRKS
HsKidins	YIAQMSQLEGGPGSTTISGRSSPHSTYYMGQSSSGGSIHSNLEQEKGDSEPKPDDGRKS
DmKidins	VASSERPSPTG-----
CeKidins	STDGSPSPMADG-----*
RnKidins	FLMKRGVDIDYSSSGVSTNEASPLDPITEDEKSDQSGSKLLPGKKSSERPSSLFQTDLKL
HsKidins	FLMKRGVDIDYSSSGVSTNDASPLDPITEDEKSDQSGSKLLPGKKSSERPSSLFQTDLKL
DmKidins	-----EMLAAVAQLOLAPIRESSEFGSPSDDQKQYGVKIS-----
CeKidins	-----FLPASVSAAPSVRFDDNINDLEREASDADSTQSRYDSK-----
RnKidins	KGGLRYQKLPSEDESGTEESDNTPLLKDDKDKKAEGKAERVCKSPESHAEPRTFIKA
HsKidins	KGGLRYQKLPSEDESGTEESDNTPLLKDDKDKKAEGKVERVPKSPESHAEPRTFIKA
DmKidins	-----N--N-----NNNNQYLHAEYNR-----SVS--S-----HS-----LQS
CeKidins	-----EN--LLEDERSAS-----P--
RnKidins	KEYLSDALLDKKDDSSDGVRSNESSPNHSLHNEAADDSOLEKANLIELEDEHSGKRGMP
HsKidins	KEYLSDALLDKKDDSSDGVRSSESSPNHSLHNEVADDSOLEKANLIELEDDSHSGKRGIP
DmKidins	LSTLVGAPVGHGGSGGS-----HLHLGNGNDLSDSTLDLMHVDVSFVGCGGGGY
CeKidins	PAHVDLMRFDSGNTSSDFVR-----SSYRTSSRYRRMSRGDLMRGMHQNDTIRSKIN
RnKidins	HSLSGLQDPPIIARMSICSEDKKSPSECSLIASSPEESWPACQKAYNLNRTPTSTVTLLNNNT
HsKidins	HSLSGLQDPPIIARMSICSEDKKSPSECSLIASSPEENWPACQKAYNLNRTPTSTVTLLNNNS
DmKidins	HRAS-----RQISISELLHSESSLPMVVVIPNASGEQAYDDTKL-----
CeKidins	FQNQ-----QTVQRRGGDDFERIQLEQFNQSDA-----
RnKidins	APTNRANQNFDEIEGIRETSQVILRPGPSPNPTAVQENLKSMAHKRSQRSSYTRLKSDA
HsKidins	APANRANQNFDEMEGIRETSQVILRPFSSPNPTTIQENLKSMTHKRSQRSSYTRLKSDP
DmKidins	-----
CeKidins	-----

Figure 12 (page 6 of 6)

SELHAASS-ESTGFGEEERESIL  
 PELHAAASSESTGFGEEERESIL  
 -----  
 -----

RnKidins  
 HsKidins  
 DmKidins  
 CeKidins

## Figure 13 (page 1 of 6)

## Human Kidins DNA Sequence

Length: 7095 July 4, 2000 16:54 Type: N

1 ggaagaatat ggccgccggg tgtggtgagg gcgacgcgct tgcagtcgcc  
51 gtctcttgc tccccgtcct ctgacatcgc ctgcagccga gcggggcccgt  
101 tccgccggag ctgaggacca ggtattcaaa taaagttaat tgcagctttc  
151 tgtgaaaATG TCAGTTTTGA TATCACAGAG CGTCATAAAT TATGTAGAGG  
201 AAGAAAACAT TCCTGCTCTG AAAGCTCTTC TTGAAAAATG CAAAGATGTA  
251 GATGAGAGAA ATGAGTGTGG CCAGACTCCA CTGATGATAG CTGCCGAACA  
301 AGGCAATCTG GAAATAGTGA AGGAATTAAT TAAGAATGGA GCTAACTGCA  
351 ATCTGGAAGA TTTGGATAAT TGGACAGCAC TTATATCTGC ATCGAAAGAA  
401 GGGCATGTGC ACATCGTAGA GGAACACTG AAATGTGGGG TTAACCTGGA  
451 GCACCGTGAT ATGGGAGGAT GGACAGCTCT TATGTGGGCA TGTTACAAAG  
501 GCCGTACTGA CGTAGTAGAG TTGCTTCTTT CTCATGGTGC CAATCCAAGT  
551 GTCACTGGTC TGCAGTACAG TGTTTACCCA ATCATTGTTGG CAGCAGGGAG  
601 AGGCCATGCA GATATAGTTC ATCTTTTACT GCAAAATGGT GCTAAAGTCA  
651 ACTGCTCTGA TAAGTATGGA ACCACCCCTT TAGTTTGGGC TGCACGAAAG  
701 GGTCATTGTTG AATGTGTGAA ACATTTATTG GCCATGGGAG CTGATGTGGA  
751 TCAAGAAGGA GCTAATTCAA TGA CTGCACT TATTGTGGCA GTGAAAGGAG  
801 GTTACACACA GTCAGTAAAA GAAATTTTGA AGAGGAATCC AAATGTAAAC  
851 TTAACAGATA AAGATGGAAA TACAGCTTTG ATGATTGCAT CAAAGGAGGG  
901 ACATACGGAG ATTGTGCAGG ATCTGCTCGA CGCTGGAACA TATGTGAACA  
951 TACCTGACAG GAGTGGGGAT ACTGTGTTGA TTGGCGCTGT CAGAGGTGGT  
1001 CATGTTGAAA TTGTTGAGC GCTTCTCCAA AAATATGCTG ATATAGACAT  
1051 TAGAGGACAG GATAATAAAA CTGCTTTGTA TTGGGCTGTT GAGAAAGGAA  
1101 ATGCAACAAT GGTGAGAGAT ATCTTACAGT GCAATCCTGA CACTGAAATA

## Figure 13 (page 2 of 6)

1151 TGCACAAAGG ATGGTGAAAC GCCACTTATA AAGGCTACCA AGATGAGAAA  
1201 CATTGAAGTG GTGGAGCTGC TGCTAGATAA AGGTGCTAAA GTGTCTGCTG  
1251 TAGATAAGAA AGGAGATACT CCCTTGCATA TTGCTATTCTG TGGAAAGGAGC  
1301 CGGAAACTGG CAGAACTGCT TTTAAGAAAT CCCAAAGATG GGCGATTACT  
1351 TTATAGGCCC AACAAAGCAG GCGAGACTCC TTATAATATT GACTGTAGCC  
1401 ATCAGAAGAG TATTTTAACT CAAATATTTG GAGCCAGACA CTTGTCTCCT  
1451 ACTGAAACAG ACGGTGACAT GCTTGGATAT GATTTATATA GCAGTGCCCT  
1501 GGCAGATATT CTCAGTGAGC CTACCATGCA GCCACCCATT TGTGTGGGGT  
1551 TATATGCACA GTGGGGAAGT GGGAAATCTT TCTTACTCAA GAAACTAGAA  
1601 GACGAAATGA AAACCTTCGC CGGACAACAG ATTGAGCCTC TCTTTCAGTT  
1651 CTCATGGCTC ATAGTGTTTC TTACCCTGCT ACTTTGTGGA GGGCTTGGTT  
1701 TATTGTTTGC CTTACGGTC CACCCAAATC TTGGAATAGC AGTGTCACTG  
1751 AGCTTCTTGG CTCTCTTATA TATATCTTTT ATTGTCATTT ACTTTGGTGG  
1801 ACGAAGAGAA GGAGAGAGTT GGAATTGGGC CTGGGTCCTC AGCACTAGAT  
1851 TGGCAAGACA TATTGGATAT TTGGAAGTCC TCCTTAAATT GATGTTTGTG  
1901 AATCCACCTG AGTTGCCAGA GCAGACTACT AAAGCTTTAC CTGTGAGGTT  
1951 TTTGTTTACA GATTACAATA GACTGTCCAG TGTAGGTGGA GAACTTCTC  
2001 TGGCTGAAAT GATTGCAACC CTCTCGGATG CTTGTGAAAG AGAGTTTGGC  
2051 TTTTGGCAA CCAGGCTTTT TCGAGTATTC AAGACTGAAG ATACTCAGGG  
2101 TAAAAAGAAA TGGAAAAAAA CATGTTGTCT CCCATCTTTT GTCATCTTCC  
2151 TTTTATCAT TGGCTGCATT ATATCTGGAA TTA CTCTTCT GGCTATATTT  
2201 AGAGTTGACC CAAAGCATCT GACTGTAAAT GCTGTCCTCA TATCAATCGC  
2251 ATCTGTAGTG GGATTGGCCT TTGTGTTGAA CTGTCGTACA TGGTGGCAAG  
2301 TGCTGGACTC GCTCCTGAAT TCCCAAAGAA AACGCCTCCA TAATGCAGCC  
2351 TCCAAACTGC ACAAATTGAA AAGTGAAGGA TTCATGAAAG TTCTTAAATG

29/79

SUBSTITUTE SHEET (RULE 26)

## Figure 13 (page 3 of 6)

2401 TGAAGTGGAA TTGATGGCCA GGATGGCAAA AACCATTGAC AGCTTCACTC  
2451 AGAATCAGAC AAGGCTGGTG GTCATCATCG ATGGATTAGA TGCCTGTGAG  
2501 CAGGACAAAG TCCTTCAGAT GCTGGACACT GTCCGAGTTC TGTTTTCAAA  
2551 AGGCCCCGTTT ATTGCCATTT TTGCAAGTGA TCCACATATT ATCATAAAGG  
2601 CAATTAACCA GAACCTCAAT AGTGTGCTTC GGGATTCAAA TATAAATGGC  
2651 CATGACTACA TGCGCAACAT AGTCCACTTG CCTGTGTTCC TTAATAGTCG  
2701 TGGACTAAGC AATGCAAGAA AATTTCTCGT AACTTCAGCA ACAAATGGAG  
2751 ACGTTCCATG CTCAGATACT ACAGGGATAC AGGAAGATGC TGACAGAAGA  
2801 GTTTCACAGA ACAGCCTTGG GGAGATGACA AAACCTGGTA GCAAGACAGC  
2851 CCTCAATAGA CGGGACACTT ACCGAAGAAG GCAGATGCAG AGGACCATCA  
2901 CTCGCCAGAT GTCCTTTGAT CTTACAAAAC TGCTGGTTAC CGAGGACTGG  
2951 TTCAGTGACA TCAGTCCCCA GACCATGAGA AGATTACTTA ATATTGTTTC  
3001 TGTGACAGGA CGATTACTGA GAGCCAATCA GATTAGTTTC AACTGGGACA  
3051 GGCTTGCTAG CTGGATCAAC CTTACTGAGC AGTGGCCATA CCGGACTTCA  
3101 TGGCTCATAT TATATTTGGA AGAGACTGAA GGTATTCCAG ATCAAATGAC  
3151 ATTAAAAACC ATCTACGAAA GAATATCAAA GAATATTCCA ACAACTAAGG  
3201 ATGTTGAGCC ACTTCTTGAA ATTGATGGAG ATATAAGAAA TTTTGAAGTG  
3251 TTTTGTCTT CAAGGACCCC AGTTCTTGTG GCTCGAGATG TAAAAGTCTT  
3301 TTTGCCATGC ACTGTAAACC TAGATCCCAA ACTACGGGAA ATTATTGCAG  
3351 ATGTTTCGTGC TGCCAGAGAG CAGATCAGTA TTGGAGGACT GCGGTACCCC  
3401 CCGCTCCCTC TACATGAGGG TCCTCCTAGG GCGCCATCAG GGTACAGCCA  
3451 GCCCCATCC GTGTGCTCTT CCACGTCCTT CAATGGGCCC TTCGCAGGTG  
3501 GAGTGGTGTC ACCACAGCCT CACAGCAGCT ATTACAGCGG CATGACGGGC  
3551 CCTCAGCATC CCTTCTACAA CAGGGGGTCA GGCCCAGCCC CAGGCCAGT  
3601 GGTATTACTG AATTCAGTGA ATGTGGATGC AGTATGTGAG AAGCTGAAAC

## Figure 13 (page 4 of 6)

3651 AAATAGAAGG GCTGGACCAG AGTATGCTGC CTCAGTATTG TACCACGATC  
3701 AAAAAGGCAA ACATAAATGG CCGTGTGTTA GCTCAGTGTA ACATTGATGA  
3751 GCTGAAGAAA GAGATGAATA TGAATTTTGG AGACTGGCAC CTTTTCAGAA  
3801 GCACAGTACT AGAAATGAGA AACGCAGAAA GCCACGTGGT CCCTGAAGAC  
3851 CCACGTTTCC TCAGTGAGAG CAGCAGTGGC CCAGCCCCGC ACGGTGAGCC  
3901 TGCTCGCCGC GCTTCCCACA ACGAGCTGCC TCACACCGAG CTCTCCAGCC  
3951 AGACGCCCTA CACACTCAAC TTCAGCTTCG AAGAGCTGAA CACGCTTGGC  
4001 CTGGATGAAG GTGCCCCCTCG TCACAGTAAT CTAAGTTGGC AGTCACAAAC  
4051 TCGCAGAACC CCAAGTCTTT CGAGTCTCAA TTCCCAGGAT TCCAGTATTG  
4101 AAATTTCAAA GCTTACTGAT AAGGTGCAGG CCGAGTATAG AGATGCCTAT  
4151 AGAGAATACA TTGCTCAGAT GTCCCAGTTA GAAGGGGGCC CCGGGTCTAC  
4201 AACCATTAGT GGCAGATCTT CTCCACATAG CACATATTAC ATGGGTCAGA  
4251 GTTCATCAGG GGGCTCTATT CATTCAAACC TAGAGCAAGA AAAGGGGAAG  
4301 GATAGTGAAC CAAAGCCCGA TGATGGGAGG AAGTCCTTTC TAATGAAGAG  
4351 GGGAGATGTT ATCGATTATT CATCATCAGG GGTTTCCACC AACGATGCTT  
4401 CCCCCCTGGA TCCTATCACT GAAGAAGATG AAAAATCAGA TCAGTCAGGC  
4451 AGTAAGCTTC TCCCAGGCAA GAAATCTTCC GAAAGGTCAA GCCTCTTCCA  
4501 GACAGATTTG AAGCTTAAGG GAAGTGGGCT GCGCTATCAA AAACCCCAA  
4551 GTGACGAGGA TGAATCTGGC ACAGAAGAAT CAGATAACAC TCCACTGCTC  
4601 AAAGATGACA AAGACAGAAA AGCCGAAGGG AAAGTAGAGA GAGTGCCGAA  
4651 GTCTCCAGAA CACAGTGCTG AGCCGATCAG AACCTTCATT AAAGCCAAAG  
4701 AGTATTTATC GGATGCGCTC CTTGACAAAA AGGATTCATC GGATTCAGGA  
4751 GTGAGATCCA GTGAAAGTTC TCCCAATCAC TCTCTGCACA ATGAAGTGGC  
4801 GGATGACTCC CAGCTTGAAA AGGCAAATCT CATAGAGCTG GAAGATGACA  
4851 GTCACAGCGG AAAGCGGGGA ATCCACATA GCCTGAGTGG CCTGCAAGAT



## Figure 13 (page 5 of 6)

4901 CCAATTATAG CTCGGATGTC CATTTGTTCA GAAGACAAGA AAAGCCCTTC  
4951 CGAATGCAGC TTGATAGCCA GCAGCCCTGA AGAAAAGTGG CCTGCATGCC  
5001 AGAAAGCCTA CAACCTGAAC CGAACTCCCA GCACCGTGAC TCTGAACAAC  
5051 AATAGTGCTC CAGCCAACAG AGCCAATCAA AATTTGATG AGATGGAGGG  
5101 AATTAGGGAG ACTTCTCAAG TCATTTTGAG GCCTAGTTCC AGTCCCAACC  
5151 CAACCACTAT TCAGAATGAG AATCTAAAAA GCATGACACA TAAGCGAAGC  
5201 CAACGTTCAA GTTACACAAG GCTCTCCAAA GATCCTCCGG AGCTCCATGC  
5251 AGCAGCCTCT TCTGAGAGCA CAGGCTTTGG AGAAGAAAGA GAAAGCATTC  
5301 TtTgagaaaa acaagcaaag gagaagagtg ttactgtacc cttatgacag  
5351 aattgtcctg gattttgact ccatccacgc ccatcacctt tctacatttt  
5401 gctgacagat aactaaccga tgatgaggcc gaggtaaaag agacatctgc  
5451 agtgtgacag aaggagcat gagaagcatg gctcaccage cagcctctgt  
5501 ggtctttgta attagaagct tcagaactca ctaatactac tgtacctttc  
5551 attggcgcat taccataa aactttttga gacgaggtga gatctgagta  
5601 taaagatagg tcagaagtat tttaaagggc ttaatgtgcc aaaaagaaaa  
5651 aaagctagag accctttttg caaacatttg gtgaccacac atttgagggg  
5701 agacgtggcg ttaggtgaag cagaagcaaa ccctgctctt aggggctcac  
5751 ctaggtgagt gcacagcctg tgacgctaca gggagaggct gagtaaaccg  
5801 agatccagcg ttctgtatgg caggggtatt gcttatcaca gaggttctga  
5851 agagtaggaa gtacataatg aagagggctt taaaaattgc caacaaagtg  
5901 agtcaccagg gctggcagta gtgtgacggg gctgtcctga gctgttagga  
5951 gagtagatgc ggggagggct ggtgacctcc gtgggtttat atgtcggaaa  
6001 ctcttctctc caaatcccag gcctggcttc cagcaccatc cagctgtgcc  
6051 caagaagcca ccctgggtctg ttctccaact cttttaaatg gtgcccaact  
6101 tttctaagtg agcttagcaa tgagaagaaa aaaaaacatg aattcttttt

## Figure 13 (page 6 of 6)

6151 ctggaaaatc agggagacat gggtaataat aggtactaat aaatatttat  
6201 agatgagtga atgaggaaat aattacatca aaaagggtcag tgacaattga  
6251 taaatgacaa ggaaatattt aattaggtaa aactaaatca ttgctctcta  
6301 tactaggata gactttatct acttcatctg ttcctaagtc agcatgntag  
6351 ttctggggaa ggatcataag aaaggaaata ctttttaaaa aaaaatttgg  
6401 aaacatgtaa caaagcaagg gtaaaatata tatatatatc tatataagtg  
6451 ctgtgactgt aaaagtgtac tttccattaa ttattagccg agttaagaga  
6501 atggtcacat tgaagtactg tgtggactag aaatgtaccc tgtcatcatg  
6551 caatgaaata ttgttatcgt ttaacatag ctcatttatg tagaatgaat  
6601 tctgggtggtt taccccaagt cacagttagg acggtagatg gtgagatcgc  
6651 agatgcgcta ttatctagat tcagtgttac attttcgatg tttatcactc  
6701 agtgggtttt tattaatatg ctgattaagt tatttactgg gccagtcatt  
6751 gtgctaaata gttgctcttt tgtgtttcat tgccttgatg tttgagtgtg  
6801 atctagcatt ttaatacagt gtttattttg catgatcttt aacaaatggt  
6851 ttaagcaatt ttaaaaaggc aggatgttat tgacattata cactgaagtc  
6901 ttaacatttt aacatttata gtgcttattt gcaaaattgt ataattagga  
6951 attatttcag agacaatggt ttctttttca ggtgagtagt tgccgcgtaa  
7001 tatcattgga gtacattctt tatactgttt gtgaaattaa tactagcata  
7051 ttaagtgtac aaatagattt agaaaacaat aaaaaattgc atgct

Figure 14a (page 1 of 2)

Rat ESTs corresponding/highly homologous to Kidins220

BLASTN 2.0.11 [Jan-20-2000]

Reference: Altschul, Stephen F., Thomas L. Madden, Alejandro A. Schaffer, Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997), "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", Nucleic Acids Res. 25:3389-3402.

Query= Kidins220 (7140 letters)

Database: embl\_est\_other; emblnew\_est\_other 1,365,225 sequences; 595,247,318 total letters

Searching.....done

Sequences producing significant alignments: (bits) Value

Score E

AI169232	AI169232	EST215067	Normalized rat kidney, Bento Soares...	1281	0.0
AI233779	AI233779	EST230467	Normalized rat lung, Bento Soares R...	1213	0.0
AW535410	AW535410	UI-R-B50-aoe-h-04-0-UI.s1	UI-R-B50 Rattus nor...	1181	0.0
AW144856	AW144856	EST292542	Normalized rat embryo, Bento Soares...	1090	0.0
AW917934	AW917934	EST349238	Rat gene index, normalized rat, nor...	1068	0.0
AI102549	AI102549	EST211838	Normalized rat embryo, Bento Soares...	1037	0.0
AW143022	AW143022	EST293317	Normalized rat brain, Bento Soares ...	955	0.0
AA900184	AA900184	UI-R-A0-bh-f-12-0-UI.s4	UI-R-A0 Rattus norveg...	904	0.0
AI385282	AI385282	UI-R-C2-mw-c-03-0-UI.s2	UI-R-C2 Rattus norveg...	866	0.0
AW917936	AW917936	EST349240	Rat gene index, normalized rat, nor...	858	0.0
AI071192	AI071192	UI-R-C2-my-h-01-0-UI.s1	UI-R-C2 Rattus norveg...	854	0.0
BE100867	BE100867	UI-R-BJ1-atx-c-03-0-UI.s1	UI-R-BJ1 Rattus nor...	833	0.0

Figure 14a (page 2 of 2)

AA998840	AA998840	UI-R-C0-ii-g-06-0-UI.s1	UI-R-C0	Rattus norveg...	805	0.0
AA998517	AA998517	UI-R-C0-ic-a-12-0-UI.s1	UI-R-C0	Rattus norveg...	781	0.0
AI717228	AI717228	UI-R-Y0-acm-f-06-0-UI.s1	UI-R-Y0	Rattus norveg...	751	0.0
BE104011	BE104011	UI-R-BX0-ari-f-07-0-UI.s1	UI-R-BX0	Rattus nor...	666	0.0
AI071379	AI071379	UI-R-C2-mw-c-03-0-UI.s1	UI-R-C2	Rattus norveg...	613	e-172
AI179761	AI179761	EST223489	Normalized rat spleen,	Bento Soares...	593	e-166
AI071701	AI071701	UI-R-C2-nc-c-08-0-UI.s1	UI-R-C2	Rattus norveg...	565	e-158
AI112518	AI112518	UI-R-Y0-mg-f-08-0-UI.s1	UI-R-Y0	Rattus norveg...	549	e-153
AA686658	AA686658	EST110947	Rat PC-12 cells,	NGF-treated (9 day...	527	e-147
AI511049	AI511049	UI-R-C2p-oc-g-12-0-UI.s1	UI-R-C2p	Rattus norv...	519	e-144
AA685317	AA685317	EST106839	Rat PC-12 cells,	untreated Rattus s...	482	e-133
H35732	H35732	EST110482	Rat PC-12 cells,	NGF-treated (9 days)	412	e-112
AA819578	AA819578	UI-R-A0-bh-f-12-0-UI.s1	UI-R-A0	Rattus norveg...	345	4e-92
AW535212	AW535212	UI-R-C4-alv-a-09-0-UI.s1	UI-R-C4	Rattus norve...	266	3e-68

Figure 14b (page 1 of 2)

Human ESTs corresponding to Kidins220

Reference: Altschul, Stephen F., Thomas L. Madden, Alejandro A. Schaffer, Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997), "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", Nucleic Acids Res. 25:3389-3402.

Query= Kidins220 masked  
(1763 letters)

Database: embl\_est\_human; emblnew\_est\_human  
2,128,571 sequences; 809,358,030 total letters

Searching.....done

Sequences producing significant alignments: (bits)			Score	E
AW954301	AW954301	EST366371 MAGE resequences, MAGC Homo sapiens...	297	2e-79
AI478209	AI478209	tm50d07.x1 NCI CGAP Kid11 Homo sapiens cDNA c...	265	4e-69
BE086866	BE086866	QV1-BT0678-300400-182-b07 BT0678 Homo sapiens...	238	3e-61
AW501669	AW501669	UI-HF-BP0p-ajh-h-05-0-UI.r1 NIH_MGC_51 Homo s...	233	1e-59
AW163610	AW163610	au96e08.y1 Schneider fetal brain 00004 Homo s...	203	2e-50
BE086801	BE086801	QV1-BT0678-290400-176-d02 BT0678 Homo sapiens...	203	2e-50
AW160324	AW160324	au66a11.y1 Schneider fetal brain 00004 Homo s...	203	2e-50
AW849765	AW849765	IL3-CT0216-240200-077-A01 CT0216 Homo sapiens...	199	2e-49
AA343860	AA343860	EST49703 Gall bladder I Homo sapiens cDNA 5' ...	199	3e-49
BE086805	BE086805	QV1-BT0678-290400-176-h06 BT0678 Homo sapiens...	195	3e-48

36/79

Figure 14b (page 2 of 2)

AA670381	AA670381	ad16g12.s1	Soares NbHFB Homo sapiens cDNA clo...	190	8e-47
M86091	M86091	EST02616	Fetal brain, Stratagene (cat#936206) Hom...	190	1e-46
T95200	T95200	ye44b05.r1	Soares fetal liver spleen 1NFLS Homo s...	111	1e-41
N77401	N77401	yz31c03.r1	Soares multiple_sclerosis_2NbHMSP Homo...	104	8e-41
Z45453	Z45453	H. sapiens partial cDNA sequence; clone c-2pg08.	147	1e-33	
T08419	T08419	EST06310	Infant Brain, Bento Soares Homo sapiens ...	141	6e-32
AW369489	AW369489	RC1-BN0005-241199-011-d07	BN0005 Homo sapiens...	138	4e-31
W37161	W37161	zb20e04.r1	Soares_fetal_lung_NbHL19W Homo sapiens...	134	1e-29
AW991382	AW991382	RC1-BN0005-060200-012-a04	BN0005 Homo sapiens...	125	5e-27
AW468986	AW468986	hc74b12.x1	NCI CGAP Gas4 Homo sapiens cDNA cl...	118	4e-25
AW991609	AW991609	RC1-BN0005-030500-015-h09	BN0005 Homo sapiens...	118	5e-25
M86121	M86121	EST02647	Fetal brain, Stratagene (cat#936206) Hom...	117	1e-24
M78657	M78657	EST00805	Hippocampus, Stratagene (cat. #936205) H...	69	2e-24
BE086873	BE086873	QV1-BT0678-300400-182-g10	BT0678 Homo sapiens...	104	7e-21
AA082204	AA082204	zn26h11.r1	Stratagene neuroepithelium NT2RAMI...	60	9e-18
N87098	N87098	L2205F	Human fetal heart, Lambda ZAP Express Homo...	68	1e-16
AW841926	AW841926	IL5-CN0024-030300-025-H06	CN0024 Homo sapiens...	85	7e-15
BE003890	BE003890	QV3-BN0097-210400-163-e08	BN0097 Homo sapiens...	83	3e-14
AW163445	AW163445	au83h06.y1	Schneider fetal brain 00004 Homo s...	77	2e-12
N83544	-	KK0817F	Human fetal heart, Lambda AP Express Homosapiens cDNA clone		

37/79

## Figure 15 (page 1 of 13)

BESTFIT of: **Human\_Kidins** check: 8556 from: 1 to: 7095to: **Rat\_Kidins** check: 131 from: 1 to: 7140

Gap Weight: 50 Average Match: 10.000  
 Length Weight: 3 Average Mismatch: -9.000

Quality: 39460 Length: 7318  
 Ratio: 5.614 Gaps: 59  
 Percent Similarity: 81.558 Percent Identity: 81.558

Match display thresholds for the alignment(s):

| = IDENTITY

: = 5

. = 1

Human Kidins x Rat Kidins

```

66 gtcctctgacatcgctgcagccgagcgggcccgttccgcccggagctgag 115
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1 gtcctctgacaaggctctgcggccctgtgagctggct.ctacagagccgag 49

116 gaccaggtattc.....aaataaagttaattgcag 145
   || || || || || || || || || || || || || || || ||
50 ggcggagtcctcagccaggccgagatacgggtgaattaaagtcggcaatag 99

146 ctttctgtgaaaATGTCAGTTTTGATATCACAGAGCGTCATAAATTATGT 195
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
100 atttctgtgaaaATGTCAGTTCTTATATCACAGAGTGTGATAAATTATGT 149

196 AGAGGAAGAAAACATTCCTGCTCTGAAAGCTCTTCTTGAAAAATGCAAAG 245
   || ||||| || ||||| ||||| ||||| ||||| ||||| |||||
150 GGAAGAAGAGAATATTCCTGCTCTGAAAGCTCTTCTTGAGAAGTGCAAAG 199

246 ATGTAGATGAGAGAAATGAGTGTGGCCAGACTCCACTGATGATAGCTGCC 295
   ||||| ||||| || ||||| ||||| ||||| ||||| ||||| |||||
200 ATGTAGATGAGAGGAACGAGTGTGGCCAGACGCCGCTGATGCTGGCTGCG 249

296 GAACAAGGCAATCTGGAATAGTGAAGGAATTAATTAAGAATGGAGCTAA 345
   || || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
250 GAGCAGGGCAATGTGGAATAGTGAAGAACTGTTAAAGAATGGAGCCAA 299

346 CTGCAATCTGGAAGATTTGGATAATTGGACAGCACTTATATCTGCATCGA 395
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
300 CTGCAATCTGGAAGATCTGGATAACTGGACTGCCCTTATATCGGCATCTA 349

```

Figure 15 (page 2 of 13)

396 AAGAAGGGCATGTGCACATCGTAGAGGAACTACTGAAATGTGGGGTTAAC 445  
| | | | | | | | | | | | | | | | | | | | | |  
350 AGGAGGGGCACATCCACATTGTGGAGGAGCTGCTGAAGAGCGGGGCCAGT 399  
446 TTGGAGCACCGTGATATGGGAGGATGGACAGCTCTTATGTGGGCATGTTA 495  
| | | | | | | | | | | | | | | | | | | | | |  
400 CTGGAGCACCGAGACATGGGTGGATGGACAGCTCTCATGTGGGCGTGCTA 449  
496 CAAAGGCCGTA CTGACGTAGTAGAGTTGCTTCTTTCTCATGGTGCCAATC 545  
| | | | | | | | | | | | | | | | | | | | | |  
450 CAAAGGCAGGACCGACGTGGTCGAGCTGCTTCTTTCTCATGGTGCCAACC 499  
546 CAAGTGTCACTGGTCTGCAGTACAGTGTTTACCCAATCATTTGGGCAGCA 595  
| | | | | | | | | | | | | | | | | | | | | |  
500 CAAGTGTCACTGGTTTGCAGTATAGTGTCTACCCAATCATCTGGGCAGCG 549  
596 GGGAGAGGCCATGCAGATATAGTTCATCTTTTACTGCAAAATGGTGCTAA 645  
| | | | | | | | | | | | | | | | | | | | | |  
550 GGCAGAGGCCATGCCGACATAGTGCATCTTCTACTGCAGAATGGTGCGAA 599  
646 AGTCAACTGCTCTGATAAGTATGGAACCAACCCCTTTAGTTTGGGCTGCAC 695  
| | | | | | | | | | | | | | | | | | | | | |  
600 AGTCAATTGCTCCGACAAGTACGGAACAACCCCTTGGTATGGGCTGCCC 649  
696 GAAAGGGTCATTTGGAATGTGTGAAACATTTATTGGCCATGGGAGCTGAT 745  
| | | | | | | | | | | | | | | | | | | | | |  
650 GAAAGGGTCATTTGGAATGTGTGAAGCACTTGTGGCCATGGGAGCTGAT 699  
746 GTGGATCAAGAAGGAGCTAATTCAATGACTGCACTTATTGTGGCAGTGAA 795  
| | | | | | | | | | | | | | | | | | | | | |  
700 GTTGATCAAGAAGGAGCCAATTTCGACGACTGCACTTATCGTGGCAGTGAA 749  
796 AGGAGGTTACACACAGTCAGTAAAAGAAATTTTGAAGAGGAATCCAAATG 845  
| | | | | | | | | | | | | | | | | | | | | |  
750 AGGAGGCTACACACAGTCCGTGAAAGAGATTTTGAAGAGGAATCCAAATG 799  
846 TAAACTTAACAGATAAAGATGGAAATACAGCTTTGATGATTGCATCAAAG 895  
| | | | | | | | | | | | | | | | | | | | | |  
800 TGAACCTAACTGACAAGGATGGAAACACAGCTCTGATGATCGCATCGAAG 849  
896 GAGGGACATACGGAGATTGTGCAGGATCTGCTCGACGCTGGAACATATGT 945  
| | | | | | | | | | | | | | | | | | | | | |  
850 GAGGGACATATTGAGATTGTGCAGGACCTGCTGGACGCCGGAACATACGT 899  
946 GAACATACCTGACAGGAGTGGGGATACTGTGTTGATTGGCGCTGTCAGAG 995  
| | | | | | | | | | | | | | | | | | | | | |  
900 GAACATACCTGACAGGAGTGGGGATACTGTATTGATCGGCGCTGTGAGAG 949



Figure 15 (page 3 of 13)

996 GTGGTCATGTTGAAATTGTTTCGAGCGCTTCTCCAAAAATATGCTGATATA 1045  
|||||  
950 GTGGTCATGTTGAAATTGTTTCGAGCACTTCTCCAAAAATATGCTGATATA 999  
1046 GACATTAGAGGACAGGATAATAAACTGCTTTGTATTGGGCTGTTGAGAA 1095  
|||||  
1000 GACATTAGAGGGCAGGACAATAAGACCGCGCTGTATTGGGCCGTTGAAAA 1049  
1096 AGGAAATGCAACAATGGTGAGAGATATCTTACAGTGCAATCCTGACACTG 1145  
|||||  
1050 GGGAAATGCGACAATGGTGAGGGATATCCTGCAGTGCAATCCGGACACGG 1099  
1146 AAATATGCACAAAGGATGGTGAAACGCCACTTATAAAGGCTACCAAGATG 1195  
|||||  
1100 AAATATGCACGAAGGATGGTGAGACACCACTGATAAAAGCCACCAAGATG 1149  
1196 AGAAACATTGAAGTGGTGGAGCTGCTGCTAGATAAAGGTGCTAAAGTGTC 1245  
|||||  
1150 AGAAATACTGAAGTGGTGGAGCTGCTGCTGGATAAGGGGGCTAAGGTGTC 1199  
1246 TGCTGTAGATAAGAAAGGAGATACTCCCTTG CATATTGCTATTCGTGGAA 1295  
|||||  
1200 TGCTGTGGACAAGAAAGGGGACACTCCCCTGCACGTTGCCATCCGAGGGA 1249  
1296 GGAGCCGGAAACTGGCAGAACTGCTTTTAAGAAATCCCAAAGATGGGCGA 1345  
|||||  
1250 GGAGCCGGAGACTGGCCGAACCTTCTTTTACGGAACCCCAAAGATGGACGG 1299  
1346 TTACTTTTATAGGCCCAACAAAGCAGGCGAGACTCCTTATAATATTGACTG 1395  
|||||  
1300 TTACTCTACAGACCCCAACAAAGCAGGCGAGACTCCCTACACATTGACTG 1349  
1396 TAGCCATCAGAAGAGTATTTTAACTCAAATATTTGGAGCCAGACACTTGT 1445  
|||||  
1350 CAGCCACCAGAAAAGCATTTTAACTCAAATATTTGGAGCCAGACACTTGT 1399  
1446 CTCCTACTGAAACAGACGGTGACATGCTTGGATATGATTTATATAGCAGT 1495  
|||||  
1400 CTCCTACGGAAACAGATGGCGACATGCTTGGTTATGATTTGTATAGCAGT 1449  
1496 GCCCTGGCAGATATTCTCAGTGAGCCTACCATGCAGCCACCCATTTGTGT 1545  
|||||  
1450 GCCCTGGCCGATATTCTCAGTGAACCCACCATGCAACCACCCATTTGTGT 1499  
1546 GGGGTTATATGCACAGTGGGGAAGTGGGAAATCTTTCTTACTCAAGAAAC 1595  
|||  
1500 GGGCTTGTATGCACAGTGGGGAAGTGGGGAAGTCTTTCTTACTCAAGAAAC 1549

1596	TAGAAGACGAAATGAAAACCTTCGCCGGACAACAGATTGAGCCTCTCTTT	1645
1550	TAGAAGATGAGATGAAGACGTTTGCAGGACAGCAGACCGAGCCCCTCTTC	1599
	·	
1646	CAGTTCTCATGGCTCATAGTGTTCCTTACCCTGCTACTTTGTGGAGGGCT	1695
1600	CAGTTCTCGTGGCTCATAGTGTTCCCTCACCCCTGCTGCTGTGTGGCGGGCT	1649
	·	
1696	TGGTTTATTGTTTGCCTTCACGGTCCACCCAAATCTTGGAATAGCAGTGT	1745
1650	CGGCCTAGTGTTTGCTTTCACGGTTGATACAAACCTCGCCATAGCCATTT	1699
	·	
1746	CACTGAGCTTCTTGGCTCTCTTATATATATTCTTTATTGTCATTTACTTT	1795
1700	CATTGAGTTTCCTGGCTCTCATATATATATTCTTCATCGTCATTTACTTC	1749
	·	
1796	GGTGGACGAAGAGAAGGAGAGAGTTGGAATTGGGCCTGGGTCCTCAGCAC	1845
1750	GGTGGACGGCGGGAAGGAGAGAGTTGGAACCTGGGCCTGGGCTCTCAGTAC	1799
	·	
1846	TAGATTGGCAAGACATATTGGATATTGGAACCTCCTCTTAAATTGATGT	1895
1800	CAGACTGGCAAGACACATTGGGTACCTGGAGCTCCTCTTCAAACCTGATGT	1849
	·	
1896	TTGTGAATCCACCTGAGTTGCCAGAGCAGACTACTAAAGCTTTACCTGTG	1945
1850	TTGTGAACCCACCTGAACTTGCGGAGCAAACCTACAAAGGCTTTACCTGTG	1899
	·	
1946	AGGTTTTTTGTTTACAGATTACAATAGACTGTCCAGTGTAGGTGGAGAAAC	1995
1900	AGGTTTTTTGTTTACAGATTACAATAGACTCTCCAGTGTGGGAGGAGAAAC	1949
	·	
1996	TTCTCTGGCTGAAATGATTGCAACCCTCTCGGATGCTTGTGAAAGAGAGT	2045
1950	TTCCTTGGCTGAAATGATCGCGACCCTCTCAGATGCTTGTGAGAGAGAAT	1999
	·	
2046	TTGGCTTTTTTGGCAACCAGGCTTTTTTCGAGTATTCAAGACTGAAGATACT	2095
2000	TCGGCTTTTTTGGCTACCCGTCTTTTTTCGAGTCCTCAGGACTGAAGAGTCT	2049
	·	
2096	CAGGGTAAAAAGAAATGGAAAAAACATGTTGTCTCCCATCTTTTGTTCAT	2145
2050	CAGGGTAAAAAGAAATGGAAAAGAACATGCTGCCTCCCTCTTTTGTTCAT	2099
	·	
2146	CTTCCTTTTTTATCATTGGCTGCATTATATCTGGAATTACTCTTCTGGCTA	2195
2100	CTTCTTTTTTATTGTTGGCTGCATTATTGCTGGGATCACCCCTTCTGGCTA	2149

2196	TATTTAGAGTTGACCCAAAGCATCTGACTGTAAATGCTGCCTCATATCA	2245
2150	TATTCAGAGTTGACCCGAAACATCTGACAGTGAATGCTATCCTCATTTCT	2199
2246	ATCGCATCTGTAGTGGGATTGGCCTTTGTGTTGAACTGTCGTACATGGTG	2295
2200	ATCGCATCTGTCTGTGGGGTTGGCCTTCGTGCTCAATTGTGCAACATGGTG	2249
2296	GCAAGTGCTGGACTCGCTCCTGAATTCCCAAAGAAAACGCCTCCATAATG	2345
2250	GCAAGTGCTGGACTCTCTCCTGAATTCTCAGAGAAAACGCCTCCATAGTG	2299
2346	CAGCCTCCAAACTGCACAAATTGAAAAGTGAAGGATTCTATGAAAGTTCTT	2395
2300	CTGCCTCCAAATTACACAAGTTAAAAAGTGAAGGATTCTATGAAAGTCCTG	2349
2396	AAATGTGAAGTGGAATTGATGGCCAGGATGGCAAAAACCATTGACAGCTT	2445
2350	AAGTGTGAGGTGGAGCTGATGGCCAGGATGGCAAAAACCATTGACAGCTT	2399
2446	CACTCAGAATCAGACAAGGCTGGTGGTCATCATCGATGGATTAGATGCCT	2495
2400	CACTCAAACCAGACAAGGCTGGTGGTTATCATTGACGGCCTGGATGCCT	2449
2496	GTGAGCAGGACAAAGTCCTTCAGATGCTGGACACTGTCCGAGTTCTGTTT	2545
2450	GCGAGCAGGATAAAGTCCTTCAGATGCTGGACACTGTCCGAGTTCTGTTT	2499
2546	TCAAAGGGCCCGTTTCATTGCCATTTTTGCAAGTGATCCACATATTATCAT	2595
2500	TCAAAGGGCCCTTTTATTGCCATTTTTGCAAGTGATCCGCATATTATCAT	2549
2596	AAAGGCAATTAACCAGAACCTCAATAGTGTGCTTCGGGATTCAAATATAA	2645
2550	AAAGGCCATCAACCAGAACCTGAATAGTGTGCTTCGTGACTCAAATATAA	2599
2646	ATGGCCATGACTACATGCGCAACATAGTCCACTTGCCTGTGTTCTTAAT	2695
2600	ACGGACATGACTACATGCGCAATATAGTTCACCTACCAGTTTTTCTTAAT	2649
2696	AGTCGTGGACTAAGCAATGCAAGAAAATTCTCGTAACTTCAGCAACAAA	2745
2650	AGTCGTGGGCTTAGCAATGCAAGGAAATTCCTTGTAACCTCAGCAACAAA	2699
2746	TGGAGACGTTCCATGCTCAGATACTACAGGGATACAGGAAGATGCTGACA	2795
2700	TGGGGACATTACATGCTCAGACACCACAGGAACACAGGAGGACACTGACA	2749

## Figure 15 (page 6 of 13)

2796 GAAGAGTTTCACAGAACAGCCTTGGGGAGATGACAAAACCTTGGTAGCAAG 2845  
|||||  
2750 GAAGAGTTTCACAGAACAGCCTTGGGGAGATGACAAAGCTTGGGAGCAAA 2799  
2846 ACAGCCCTCAATAGACGGGACACTTACCGAAGAAGGCAGATGCAGAGGAC 2895  
|| || ||||| ||| |||||  
2800 ACGGCTCTCAACAGAAGGGACACTTACCGCAGAAGACAGATGCAGAGAAC 2849  
2896 CATCACTCGCCAGATGTCCTTTGATCTTACAAAACCTGCTGGTTACCGAGG 2945  
||||| | |||||  
2850 CATCACGAGGCAGATGTCCTTTGATCTCACAAAGCTGCTGGTCACTGAGG 2899  
2946 ACTGGTTCAGTGACATCAGTCCCCAGACCATGAGAAGATTACTTAATATT 2995  
| |||||  
2900 ATTGGTTCAGTGACATCAGCCCTCAGACCATGAGACGGCTACTCAATATT 2949  
2996 GTTCTGTGACAGGACGATTACTGAGAGCCAATCAGATTAGTTTCAACTG 3045  
||||| || |||||  
2950 GTTCTGTGACAGGGCGGCTACTAAGAGCCAATCAGATTACTTTCAACTG 2999  
3046 GGACAGGCTTGCTAGCTGGATCAACCTTACTGAGCAGTGGCCATACCGGA 3095  
||||| || |||||  
3000 GGACAGGCTAGCCAGCTGGATCAACCTCACCGAGCAGTGGCCGTACCGGA 3049  
3096 CTTTCATGGCTCATATTATTTGGAAGAGACTGAAGGTATTCCAGATCAA 3145  
| || ||||| ||| |||||  
3050 CGTCTTGGCTTATACTATATCTGGAAGAGACTGAAGGTCTCCCTGATCAG 3099  
3146 ATGACATTAAAAACCATCTACGAAAGAATATCAAAGAATATTCCAACAAC 3195  
||||| ||||| ||||| ||| |||||  
3100 ATGACCTTAAAGACCATCTATGAGAGAATATCGAAGAATATCCCAACAAC 3149  
3196 TAAGGATGTTGAGCCACTTCTTGAAATTGATGGAGATATAAGAAATTTTG 3245  
||| |||||  
3150 TAAAGATGTTGAGCCGCTTCTTGAAATTGATGGAGATATAAGAACTTTG 3199  
3246 AAGTGTTTTTGTCTTCAAGGACCCAGTTCTTGTGGCTCGAGATGTAAAA 3295  
|||| || |||||  
3200 AAGTCTTCTTGTCTTCAAGGACTCCAGTTCTCGTGGCTCGAGATGTAAAG 3249  
3296 GTCTTTTGGCCATGCACTGTAAACCTAGATCCCAAACCTACGGGAAATTAT 3345  
||||| ||||| |||||  
3250 ACCTTTTGGCGTGACAGTAAACCTGGATCCCAAACCTTCGGGAGACCAT 3299  
3346 TGCAGATGTTGCTGCTGCCAGAGAGCAGATCAGTATTGGAGGACTGGCGT 3395  
||||| || |||||  
3300 CGCAGATGTCCGAGCTGCAAGAGAGCAGATCAACATAGGAGGCCTGGCCT 3349

3396 ACCCCCCGCTCCCTCTACATGAGGGTCCTCCTAGGGCGCCATCAGGGTAC 3445  
| | | | | | | | | | | | | | | | | | | | | | | | | | | |  
3350 ATCCCCCGCTCCCTCTGCACGAAGGCCCTCCTCGGCCACCTTCTGGGTAC 3399  
| | | | | | | | | | | | | | | | | | | | | | | | | | | |  
3446 AGCCAGCCCCCATCCGTGTGCTCTTCCACGTCTTCAATGGGCCCTTCGC 3495  
| | | | | | | | | | | | | | | | | | | | | | | | | | | |  
3400 AGTCAGCCTGCATCTGTCTGCTCTTCAGCATCTTTCAACGGGCCGTTCCC 3449  
| | | | | | | | | | | | | | | | | | | | | | | | | | | |  
3496 AGGTGGAGTGGTGTCAACCACAGCCTCACAGCAGCTATTACAGCGGCATGA 3545  
| | | | | | | | | | | | | | | | | | | | | | | | | | | |  
3450 AGGCGGGGTTGTGTCTCCACAGCCCCACAGCAGTTACTACAGCGGCTTGT 3499  
| | | | | | | | | | | | | | | | | | | | | | | | | | | |  
3546 CGGGCCCTCAGCATCCCTTCTACAACAGG..... 3574  
| | | | | | | | | | | | | | | | | | | | | | | | | | | |  
3500 CCGGACCCCAGCACCCCTTCTACAACAGGCCATTCTTTGCCCCATACCTT 3549  
.  
.  
  
3575 .....GGGT.....CAGGCCCAGCCCCAGGCC 3597  
| | | | | | | | | | | | | | | | | | | | | | | | | | | |  
3650 ACTCTGGGTTTAACAAACAGAGACAGGGCAGCCGTCCCTGCCACAGGCAG 3699  
| | | | | | | | | | | | | | | | | | | | | | | | | | | |  
3598 AGTGGTATTACTGAATTCAGTGAATGTGGATGCAGTATGTGAGAAGCTGA 3647  
| | | | | | | | | | | | | | | | | | | | | | | | | | | |  
3700 CTCAGTGTTACTGAGTTCAATGACCGTGGATGTCGT.TGTGAGAACTGA 3748  
| | | | | | | | | | | | | | | | | | | | | | | | | | | |  
3648 AACAAATAGAAGGGCTGGACCAGAGTATGCTGCCTCAGTATTGTACCACG 3697  
| | | | | | | | | | | | | | | | | | | | | | | | | | | |  
3749 GACAGATAGAAGGGCTGGACCAGAGCATGATGCCTCAGTACTGCACAACA 3798  
| | | | | | | | | | | | | | | | | | | | | | | | | | | |  
3698 ATCAAAAAGGCAAACATAAATGGCCGTGTGTTAGCTCAGTGTAACATTGA 3747  
| | | | | | | | | | | | | | | | | | | | | | | | | | | |  
3799 ATCAAAAAGGCGAACATAAATGGCCGGGTATTGTCTCAGTGTAACATTGA 3848  
| | | | | | | | | | | | | | | | | | | | | | | | | | | |  
3748 TGAGCTGAAGAAAGAGATGAATATGAATTTTGGAGACTGGCACCTTTTCA 3797  
| | | | | | | | | | | | | | | | | | | | | | | | | | | |  
3849 TGAGCTGAAGAAAGAGATGGCTATGAACTTTGGAGACTGGCATCTCTTCA 3898  
| | | | | | | | | | | | | | | | | | | | | | | | | | | |  
3798 GAAGCACAGTACTAGAAATGAGAAACGCAGAAAGCCACGTGGTCCCTGAA 3847  
| | | | | | | | | | | | | | | | | | | | | | | | | | | |  
3899 GAAGCATGGTCTTAGAAATGAGAAAGTGTGGAAAGCCAGGTGGTCCCTGAA 3948  
| | | | | | | | | | | | | | | | | | | | | | | | | | | |  
3848 GACCCACGTTTCCTCAGTGAGAGCAGCAGTGGCCCAGCCCCGCACGGTGA 3897  
| | | | | | | | | | | | | | | | | | | | | | | | | | | |  
3949 GACCCTCGTTTTTCTCAATGAAAACCTCCAGTGCCCCAGTTCTCATGGAGA 3998  
| | | | | | | | | | | | | | | | | | | | | | | | | | | |  
3898 GCCTGCTCGCCGCGCTTCCCACAACGAGCTGCCTCACACCGAGCTCTCCA 3947  
| | | | | | | | | | | | | | | | | | | | | | | | | | | |  
3999 GTCTGCTCGCCGCCCTTCTCACACTGAGCTGCCCTCACGGAGCTTTCCA 4048

3948	GCCAGACGCCCTACACACTCAACTTCAGCTTCGAAGAGCTGAACACGCTT	3997
4049	GCCAGACTCCATACACCCTGAACTTCAGCTTTGAAGAACTGAACACGCTT	4098
	· · · · ·	
3998	GGCCTGGATGAAGGTGCCCCTCGTCACAGTAATCTAAGTTGGCAGTCACA	4047
4099	GGCCTTGATGAGGGTGCCCCTCGTCACAGTAATCTAAGTTGGCAGTCACA	4148
	· · · · ·	
4048	AACTCGCAGAACCCCAAGTCTTTCGAGTCTCAATTCCCAGGATTCCAGTA	4097
4149	AACTCGCAGAACCCCAAGTCTCTCGAGTCTCAATTCCCAGGACTCCAGTA	4198
	· · · · ·	
4098	TTGAAATTTCAAAGCTTACTGATAAGGTGCAGGCCGAGTATAGAGATGCC	4147
4199	TTGAAATTTCAAAGCTTACTGATAAGGTGCAGGCCGAGTATAGAGACGCC	4248
	· · · · ·	
4148	TATAGAGAATACATTGCTCAGATGTCCCAGTTAGAAGGGGGGGCCCCGGGTC	4197
4249	TATAGAGAGTACATTGCTCAGATGTCCCAGTTAGAAGGGGGGCACAGGGTC	4298
	· · · · ·	
4198	TACAACCATTAGTGGCAGATCTTCTCCACATAGCACATATTACATGGGTC	4247
4299	GTCCACAATAAGCGGCAGATCTTCTCCACATAGCACATACTACATAGGCC	4348
	· · · · ·	
4248	AGAGTTCATCAGGGGGCTCTATTCATTCAAACCTAGAGCAAGAAAAGGGG	4297
4349	AGAGCTCTTCAGGGGGCTCCATCCATTCTACTCTAGAACAGGAAAGGGGG	4398
	· · · · ·	
4298	AAGGATAGTGAACCAAAGCCCGATGATGGGAGGAAGTCCTTTCTAATGAA	4347
4399	AAGGAGGGTGAGCTAAAGCAGGAGGATGGGCGGAAGTCATTTTTAATGAA	4448
	· · · · ·	
4348	GAGGGGAGATGTTATCGATTATTCATCATCAGGGGTTTCCACCAACGATG	4397
4449	GAGGGGGGATGTCATAGATTACTCCTCATCAGGGGTGTCCACTAATGAGG	4498
	· · · · ·	
4398	CTTCCCCCTGGATCCTATCACTGAAGAAGATGAAAAATCAGATCAGTCA	4447
4499	CCTCACCCCTGGACCCCATTAAGTGAAGAAGATGAGAAATCCGACCAGTCA	4548
	· · · · ·	
4448	GGCAGTAAGCTTCTCCAGGCAAGAAATCTTCCGAAAGGTCAAGCCTCTT	4497
4549	GGTAGTAAGCTACTCCAGGCAAGAAATCCTCAGAACGGCCCAGTCTCTT	4598
	· · · · ·	
4498	CCAGACAGATTTGAAGCTTAAGGGAAGTGGGCTGCGCTATCAAAAACCTCC	4547
4599	CCAGACAGACTTGAAGCTTAAGGGGGGTGGTCTTCGCTACCAGAAACTGC	4648

4548	CAAGTGACGAGGATGAATCTGGCACAGAAGAATCAGATAACACTCCACTG	4597
4649	CCAGTGATGAAGATGAGTCGGGGACAGAAGAGTCAGATAACACCCCCTG	4698
	.	
4598	CTCAAAGATGACAAAGACAGAAAAGCCGAAGGGGAAAGTAGAGAGAGTGCC	4647
4699	CTCAAAGATGATAAGGACAAAAAGGCTGAAGGCCAAAGCAGAGAGAGTGTG	4748
	.	
4648	GAAGTCTCCAGAACACAGTGCTGAGCCGATCAGAACCTTCATTAAAGCCA	4697
4749	CAAGTCCCCAGAGCACAGTGCCGAGCCCATCAGAACGTTTCATTAAAGCAA	4798
	.	
4698	AAGAGTATTTATCGGATGCGCTCCTTGACAAAAAGGATTTCATCGGATTCA	4747
4799	AAGAGTATTTGTCCGATGCCCTCCTTGACAAAAAGGATTTCCTCAGATTGC	4848
	.	
4748	GGAGTGAGATCCAGTGAAAGTTCTCCCAATCACTCTCTGCACAATGAAGT	4797
4849	GGAGTGAGGTCCAATGAGAGTTCTCCCAACCACTCTCTCCACAATGAGGC	4898
	.	
4798	GGCGGATGACTCCCAGCTTGAAAAGGCCAATCTCATAGAGCTGGAAGATG	4847
4899	AGCAGACGACTCCCAACTGGAAAAGGCCAACCCTCATAGAGCTTGAAGATG	4948
	.	
4848	ACAGTCACAGCGGAAAGCGGGGAATCCACATAGCCTGAGTGGCCTGCAA	4897
4949	AGGGCCACAGTGGGAAGCGGGGGATGCCACACAGTCTGAGTGGCTTGCAA	4998
	.	
4898	GATCCAATTATAGCTCGGATGTCCATTTGTTTCAGAAGACAAGAAAAGCCC	4947
4999	GATCCAATTATAGCTCGGATGTCCATTTGCTCGGAAGACAAGAAAAGCCC	5048
	.	
4948	TTCCGAATGCAGCTTGATAGCCAGCAGCCCTGAAGAAACTGGCCTGCAT	4997
5049	TTCCGAGTGCAGCCTGATTGCTAGCAGCCCTGAAGAAAGCTGGCCTGCGT	5098
	.	
4998	GCCAGAAAGCCTACAACCTGAACCGAACTCCCAGCACCGTGACTCTGAAC	5047
5099	GCCAGAAGGCCTACAATCTAAACCGAACGCCCAGCACTGTGACTCTGAAC	5148
	.	
5048	AACAATAGTGCTCCAGCCAACAGAGCCAATCAAAATTTTCGATGAGATGGA	5097
5149	AACAACACTGCACCCACCAACAGAGCCAATCAGAACTTCGATGAGATAGA	5198
	.	
5098	GGGAATTAGGGAGACTTCTCAAGTCATTTTGAGGCCTAGTTCCAGTCCCA	5147
5199	GGGAATCAGGGAGACGTCCCGAGGTCATTCTGAGGCCTGGCCCCAGTCCCA	5248

Figure 15 (page 10 of 13)

5148 ACCCAACCACTATTCAGAATGAGAATCTAAAAAGCATGACACATAAGCGA 5197  
||||||| || | ||||||||| || || | ||||||| | || |||||||  
5249 ACCCAACTGCTGTGCAGAATGAAAACCTGAAGAGCATGGCGCACAAAGCGA 5298  
5198 AGCCAACGTTCAAGTTACACAAGGCTCTCCAAAGATCCTCCGGAGCTCCA 5247  
||||| || || ||||||||||||||||| |||||| | | |||||| ||  
5299 AGCCAGCGCTCCAGTTACACAAGGCTCTCGAAAGATGCATCTGAGCTACA 5348  
5248 TGCAGCAGCCTCTTCTGAGAGCACAGGCTTTGGAGAAGAAAGAGAAAGCA 5297  
| |||||||| || ||||||||||||||||||||||||| ||||| |||||  
5349 T...GCAGCCTCCTCGGAGAGCACAGGCTTTGGAGAAGAGAGAGAGAGCA 5395  
5298 TTCTTtgagaaaaacaagcaaaggagaagagtgttactgtacccttatga 5347  
|||| |||| ||||||||| ||||| |||||  
5396 TTCTCtgag.....gaaggagaag.....tga 5417  
5348 cagaattgtcctggattttgactccatccacgcccattcacctttctacat 5397  
| | ||| || ||||| | | | | ||| || || |  
5418 aaagagtgtgacagactttgagttgaccattcctcattactcatcatgtt 5467  
5398 tttgc.tgacagataactaaccgatgatgaggccgaggtaaaagagacat 5446  
| ||| | ||||| | || | | ||| |||| || | ||  
5468 tatgcatcacagacacctgataggtggtcaggcactggaagcag..... 5511  
5447 ctgcagtgtgacagaagggagcatgagaagcatggctcaccagccagcct 5496  
| | | ||| ||| ||||| | ||||||| || |  
5512 .....acaggaaagcctga.aagcacag.tcaccagtcatttt 5547  
5497 ctgtggtctttgtaattagaagc...ttcagaactcactaataactactgt 5543  
| | ||| || || | || | || | ||||| ||||| |  
5548 c.attgtcaccgtcatcccaggcactttaaggtctcactgccactattac 5596  
5544 acctttcattggcgcattaccccataaaaactttttgagacgaggtgagat 5593  
||||||| |||| || | | | ||||| |||  
5597 acctttc.ctggcacacagctaggcagagc..gatgagagcagg..... 5637  
5594 ctgagtataaagataggtcagaagt.....attttaagggcttaat 5635  
|||| | ||| ||||||||| |||| | || |||||||||  
5638 ctgaagacaaa.ataggtcaaaagtgtattttaagttgaaagggcttaa. 5685  
5636 gtgccaaaaaagaaaaaaagctagagaccctttttgcaaacatttggtgac 5685  
|| || ||||||||| |||||| | ||| |||| || ||  
5686 ..atcacaaaggaaaaaaacttagagatctcctttataaaactattcatgta 5733  
5686 cacacatttgaggggaagacgtggcgttaggtgaagcagaagcaaaccctg 5735  
| ||||||||||||| | ||||| ||||||||| | || |  
5734 c...catttgaggggaagatacagtgttagatgaagcagaaaccaatttta 5780



## Figure 15 (page 11 of 13)

5736 ctcttaggggctcacctaggtg.....agtgcacagcctgtg 5772  
| | | | | | | | | | | | | | | | | | | | | |  
5781 ttgttagtggttaatcttggtgcagatttataagtttttagagtagcccg 5830  
5773 acgctacagggagaggctgagtaaaccgagatccagcgttctgtatggca 5822  
| | | | | | | | | | | | | | | | | | | | | |  
5831 agactaaaagtgaatacttagcaaatggatagccagtgttctatatag.. 5878  
5823 ggggtattgcttatcacagaggttctgaag....agtaggaagtacataa 5868  
| | | | | | | | | | | | | | | | | | | | | |  
5879 gaatcattgctttt..cagagggcttaaagtttaagtagaaaatatatac 5926  
5869 tgaagagggcttttaaaaattgccacaagaagtgagtcaccaggggctggcag 5918  
| | | | | | | | | | | | | | | | | | | | | |  
5927 tcaagaaggcgataaaaagctgatgagaaagtgagttagcagaacccaaag 5976  
5919 tagtgtgacggggctgtcctgagctgtaggagagtagatgcgggggaggg 5968  
| | | | | | | | | | | | | | | | | | | | | |  
5977 ccgtg...ctgggccgcggtgactcattagcagag..gaggaggggaggg 6021  
5969 ctggtgacctccgtggggtttatatgtcggaaactcttctctccaaatccc 6018  
| | | | | | | | | | | | | | | | | | | | | |  
6022 c.....agtatattcctgggatac...tctctccaga..cc 6052  
6019 aggcctggcttccagcaccatccagctgtgcccaagaagccaccctgggtc 6068  
| | | | | | | | | | | | | | | | | | | | | |  
6053 cagcctggcttctgacatcatccacctgtgccctcaaaaccgtcttagtc 6102  
6069 tgttctccaactcttttaaatgggtgcccaacttttctaagtgagcttagc 6118  
| | | | | | | | | | | | | | | | | | | | | |  
6103 tgttctgcaactc..ttaagtg.....acatacctaactcagctcatg 6143  
6119 aatgagaagaaaaa..... 6133  
| | | | | | | | | | | | | | | | | | | | | |  
6144 gctaagggaataaataaagtgtcctgggtgattaaactctggacctcc 6193  
6134 .....aaacatgaat 6143  
| | | | | | | | | | | | | | | | | | | | | |  
6194 cacatctaagtcgagttgacaaactgcatccccagcccaaacatctgt 6243  
6144 tctt.....tttctggaaaatcaggagacatgggtaataataggtacta 6188  
| | | | | | | | | | | | | | | | | | | | | |  
6244 tcttccagagatctggaaagacatgggcataagagt.....cca 6282  
6189 ataaatatattatagatgagtgaatgaggaaataattacatcaaaaaggt. 6237  
| | | | | | | | | | | | | | | | | | | | | |  
6283 gtaaactctta...atgaaggtgtcgggaagtgttcccttaaagggggt 6329  
| | | | | | | | | | | | | | | | | | | | | |

6238 .....cagtgacaattgataaatgacaaggaaatatttaatttaggtaa... 6282  
| | | | | | | | | | | | | | | | | | | | | |  
6330 agtgacagtgcactagacaagtgc.....agaatttagtttgcctgaca 6374  
ctaaatcattgctctctatactaggatagacttttatctacttcattctgtt 6332  
| | | | | | | | | | | | | | | | | | | | | |  
6375 ttaaactcactgccttctgggccgaatag..tta.....gtatctgtt 6416  
cctaagtcagcatgttagttctggggaaggatcataagaaaggaaatact 6382  
| | | | | | | | | | | | | | | | | | | | | |  
6417 cctaaatcagcatttgagttctgcgggggttacaaaatggaggca..tgg 6464  
ttttaaaaaaaaatttggaacatgtaacaaagcaagggtaaaatatata 6432  
| | | | | | | | | | | | | | | | | | | | | |  
6465 ttttaagaaggaatttgaaagcatccagggaagcaagaacagagt..... 6509  
tatatatctatataagtgctgtgactgtaaaagtgtactttccattaatt 6482  
| | | | | | | | | | | | | | | | | | | | | |  
6510 .....cgagctggtgtgaccatga.....cacttcccatt.att 6543  
attagccgagttaagagaatggc....acattgaagtactgtgtggact 6528  
| | | | | | | | | | | | | | | | | | | | | |  
6544 gttagcaaagctaggagaatgaccacatacatttaagtatcatgtaggct 6593  
aga.aatgtaccctgtcatcatgcaatgaaatattgttatcgttttaaca 6577  
| | | | | | | | | | | | | | | | | | | | | |  
6594 agagcgtgcctgtcatcatcatgcattgagaccgt.ttgtcctttgatta 6642  
tagctcatttatgtagaatgaattctgggtggtttacccaagtacagtt 6627  
| | | | | | | | | | | | | | | | | | | | | |  
6643 tagtgtgtttatgttag.accagtcctagcagtatgtcctgagccatgggc 6691  
aggacggtagatgggtgagatgcgagatgcgctattatctagattcagtg 6677  
| | | | | | | | | | | | | | | | | | | | | |  
6692 aggattgtagatggcgaggtggc.....aggttttagtgc 6725  
tacatttttcgatgtttatcactcagtgggtttttattaatatgctgatta 6727  
| | | | | | | | | | | | | | | | | | | | | |  
6726 t..cgttgagatgtccatctctgactggacctgt..... 6757  
agttattttactgggccagtcattgtgctaaatagttgctcttttggtgtt 6777  
| | | | | | | | | | | | | | | | | | | | | |  
6758 .gtggtttacttg.....tagctgctc...cgtgaaa 6785  
cattgccttgatgtttgagtgtaatctagcatttttaatacagtggtttatt 6827  
| | | | | | | | | | | | | | | | | | | | | |  
6786 cactgccttgctgtttgagtgtaattgagcattttcatgaagtgtatatq 6835

```
6828 ttgcatgatctttaacaaatgttttaagcaatttttaaaaaggcaggatgt 6877  
      |||||  
6836 ttgcatgatctttaacaggtatttttag.acattaagaacagcatgatata 6884  
      . . . . .  
6878 tattgacattatacactgaagtcttaacatttttaacatttatagtgccta 6927  
      || ||||| | ||||| |||  
6885 tactgaca.....tatgcattagtag..... 6905  
      . . . . .  
6928 tttgcaaaattgtataattaggaattatttcagagacaatgt.tttcttt 6976  
      | |||| | | |||| | || ||||| || | | || ||  
6906 tatgcagtatcatgtaatggg....tacttcagaggcagtttggtttctt 6951  
      . . . . .  
6977 ttcaggtgagtagttgccgcgtaatatcattggagtacattctttatact 7026  
      | ||||| ||||| || | |||| || | | || | |||  
6952 tccaggtgagtagttatagctcacagtcatgggaatcctttattcacact 7001  
      . . . . .  
7027 gtttgtgaaattaatactagcatattaagtgtacaaatagatttagaaaa 7076  
      || ||||| | || || | | || || ||| || | ||  
7002 gtctgtgaagtcgatgctggagtgccaaatgcacacggagactagg..aa 7049  
      .  
7077 caataaaaaattgcatgc 7094  
      ||||| | |||||  
7050 caataaagacttgcatgc 7067
```

Figure 16

CCCAATGGT	ACTTAACACT	CAAGGCCCTT	CAGAGGCTGA	CCCCTGCTCC	ATCTACTGCC	34620
CCCTCCCATG	CTGCCACCCC	AGCCCCAAGG	CTGGCAAATA	TTGACAGTTC	CTGGAACTTT	34680
GGGCCACCCCT	CCTTCTCATC	TCTATTCTTT	TCCAAAGTGT	TGTTCCCCGA	CCTGGACATT	34740
CTTCCTAGCC	CCATCCAACT	GATGTGATGG	ACAAACTATA	CTTCAGATAC	AGCTGGAGGT	34800
GCGGGGGGG	GGGGTCCCTT	CTTTTGTGAA	CCTTCCTAGT	GGACCCCAAG	CAGAGTTGGT	34860
AACTCCATCC	TCTGTGCTTA	CCTCCTTACT	GCACGTGACC	TAGGACATGG	TCAGGACTCA	34920
GAAAGGCTGT	GCTAAATAAC	GCACATTAA	CCTGTTTCCT	TGTCTTCTCC	CCTACTCTAA	34980
GGTTGGCTTC	ATCCTTATAT	CCTACTGATA	TGGTTTGGCT	CTGTGACCCC	ACCCAATCT	35040
CATCTCGAAT	TGTATTCCCC	ATAATCCCCA	CGTGTCAAGG	GAGGGACTTG	GTCGGGGGTG	35100
ATTGGATCAC	GTCAGCGGTT	TCCCCCGTGC	TGTTCTCGTG	ATAGTGAGTG	AATTCTCAGG	35160
AGATCTGTTG	GATGTAAAG	GGGCTTTTGA	CCCTTTTGCT	GGGCACCTCT	CTTCCCTGCAG	35220
CCATGTGACG	AAGGACGTGT	TTGCTTCCCT	TTCCGCCATG	ATTTTAAAGT	TCCTGAGGCT	35280
TCTCCAGCCA	TGCAGAACTG	TGAGTCAATT	AAACCTCTTT	TCTTTATAAA	TTACCCCTGTC	35340
TCAGGTGGTT	CTTTATAGCA	GTGTGAAAAC	GGGCTAATAC	ACCTACCATA	CTTCTAACCC	35400
CCAACGCCCTG	ACCCACCCCTG	CCCTGTTTAC	AACAGTGGCT	GGCCTTAGCA	AGTCATTATC	35460
GAAGAAATCA	CACCATTTAG	GACTCCACTG	TACACGTTGG	GGTGTTCCGT	CCTTGTTTCA	35520
TGAGGGGCAC	ACCAACTTCC	GTTTTTGTTT	GTTTGTTTGT	TTGTTTGGAG	ATAGGATCTT	35580
GCCCTGTTGC	CCTGGCTGAA	GTGCAGTGGT	GTGATCTCAG	CTCACTGCAG	CCTCGAACTC	35640
CTGGGCTCAA	GTGATCCTCT	CACCTCTGTC	TCCTAAGTGG	CTGGGACTAC	AGGTGTGCAA	35700
CACCACGCCCT	GGCTAAGTTT	TGTATTTTGG	TATTTTTTTT	TTTTTTTTTT	TTTTTTGTAG	35760
AGAAGAGGTC	TCTCCATGTT	GCCCAGGCTG	GTTTCAAAC	CCTGAGCTCA	AGTGATCCTC	35820
CGTCTTGGCC	TCCCAAATG	CTGGGATTAC	AGGAGTGAGC	CATTGTCCCC	ACCCCGCACC	35880
AACCTTCTGAT	GAAGTCAGGC	TCCTTATTGG	CTTTTTCCTG	AGTTGCCAGG	TTGACACATT	35940
GGTGGGTGCT	TTTCGATGGT	TCGCAACCTA	CGCACGTCGT	AATCATCCTT	GGACCTTTTC	36000
AAAAATCAAG	ATTGACGCTT	GTTCGCAGCA	AGGCCTAGGA	ATGTGGAGGC	TGAAGGAGGT	36060
GAGGGGAGGT	GCTGGGAGAC	CTGGAAGGAG	TGTGAACACT	AGCCCCCGCG	AAAGGCAGAG	36120
GTCTCGCTCC	TGGATCCGGA	AGAAGGCTGC	TGAGCGGGCT	TTCAATTTCTT	ATGTATGTGC	36180
AGTTTCTTCT	GAAGATATAA	TTGATGTTTA	TTTCTCTGAA	TAAGCTTGAT	TATAATGGAA	36240
AATAGGGATA	ATAAATTAAC	TGTAATAAGA	GTGACATTGA	CACGTAAAAA	CCTTGCCTGG	36300
GGCGCTTGTC	CTGGGTGTCC	CCGCCCCGCG	GCCCCTCCAG	CCCCGCCCTG	TCTACGCCCC	36360
GCCCCGCCCC	TCCCCCTCCAC	GCCCAGGGGG	CGTGGCAACC	CGGAGTCCCT	CCCCTACCCA	36420
GGGCACGTGA	TCTGTCAAGT	CGCGGAGAGG	GAAGAATATG	GCCGCCGGGT	GTGGTGAGGG	36480
CGACGCGCTT	GCAGTCGCCG	TCTCTTGCTT	CCCCGTCTCT	TGACATCGCC	TGCAGCCGAG	36540
CGGGCCCGTT	CGCCCGGAGC	TGAGGACCAG	GT			36600

Figure 17 (page 1 of 3)

1	gtcctctgac	aaggtctgcg	gccctgtgag	ctggctctac	agagccgagg	gcggagtcct
61	cagccaggcc	gagatacgg	gaattaaagt	cggcaataga	tttctgtgaa	aatgtcagtt
121	cttatatcac	agagtgtgat	aaattatgtg	gaagaagaga	atattcctgc	tctgaaagct
181	cttcttgaga	agtgcacaga	tgtagatgag	aggaacgagt	gtggccagac	gccgctgatg
241	ctggctgagg	agcagggcaa	tgtggaaata	gtgaaagaac	tgttaaagaa	tggagccaac
301	tgcaatctgg	aagatctgga	taactggact	gcccttatat	cggcatctaa	ggaggggac
361	atccacattg	tggaggagct	gctgaagagc	ggggccagtc	tggagcaccg	agacatgggt
421	ggatggacag	ctctcatgtg	ggcgtgctac	aaaggcagga	cgcagctggg	cgagctgctt
481	ctttctcatg	gtgccaaccc	aagtgtcact	ggtttgtata	gtgtctaccc	aatcatctgg
541	gcagcggggc	gaggccatgc	cgacatagt	catcttctac	tgcagaatgg	tgcgaaagtc
601	aattgctccg	acaagtacgg	aacaaccccc	ttgggtatggg	ctgcccgaag	gggtcatttg
661	gaatgtgtga	agcacttggt	ggccatggga	gctgatgttg	atcaagaagg	agccaattcg
721	atgactgcac	ttatcgtggc	agtgaaggga	ggctacacac	agtccgtgaa	agagatcttg
781	aaaagggaatc	caaatgtgaa	cctaactgac	aaggatggaa	acacagctct	gatgatcgca
841	tcgaaggagg	gacatattga	gattgtgcag	gacctgctgg	acgccggaac	atacgtgaac
901	atacctgaca	ggagtgggga	tactgtattg	atcggcgctg	tgagagggtg	tcatgttgaa
961	attgttcgag	cacttctcca	aaaatatgct	gatatagaca	ttagagggca	ggacaataag
1021	accgcgctgt	attgggcccgt	tgaaggggga	aatgcgacaa	tgggtgaggga	tatcctgcag
1081	tgcaatccgg	acacggaaat	atgcacgaag	gatggtgaga	caccactgat	aaaagccacc
1141	aagatgagaa	atattgaagt	ggtggagctg	ctgctggata	aggggggcta	ggtgtctgct
1201	gtggacaaga	aaggggacac	tcccctgcac	gttgccatcc	gagggaggag	ccggagactg
1261	gccgaacttc	ttttacggaa	ccccaaagat	ggacggttac	tctacagacc	caacaaagca
1321	ggcgagactc	cctacaacat	tgactgcagc	caccagaaaa	gcatttttaac	tcaaataattt
1381	ggagccagac	acttgtctcc	tacggaaaca	gatggcgaca	tgcttggtta	tgatttgtat
1441	agcagtcccc	tggccgatat	tctcagtga	cccaccatgc	aaccacccat	ttgtgtgggc
1501	ttgtatgcac	agtggggaag	tgggaagtct	ttcttactca	agaaactaga	agatgagatg
1561	aagacgtttg	caggacagca	gaccgagccc	ctcttccagt	tctcgtggct	catagtgttc
1621	ctcaccctgc	tgctgtgtgg	cgggctcggc	ctagtgtttg	ctttcacggt	tgatacaaac
1681	ctcgccatag	ccatttcatt	gagtttcctg	gctctcatat	atatattctt	catcgtcatt
1741	tacttcggtg	gacggcgggg	aggagagagt	tggaaactggg	cctgggctct	cagtaccaga
1801	ctggcaagac	acattgggta	cctggagctc	ctcttcaaac	tgatgtttgt	gaaccacact
1861	gaactgccgg	agcaaaactac	aaaggcttta	cctgtgaggt	ttttgtttac	agattacaat
1921	agactctcca	gtgtgggagg	agaaacttcc	ttggctgaaa	tgatcgcgac	cctctcagat
1981	gcttgtgaga	gagaattcgg	ctttttggct	acccgtcttt	ttcgagtctt	caggactgaa
2041	gagtctcagg	gtaaaaagaa	atggaaaaaa	acatgctgcc	tcccctcttt	tgtcatcttt
2101	ctttttattg	ttggctgcat	tattgctggg	atcacccttc	tggctatatt	cagagttgac
2161	ccgaaacatc	tgacagtga	tgctatcctc	atttctatcg	catctgtcgt	ggggttggcc
2221	ttcgtgctca	attgtcgaac	atgggtggca	gtgctggact	ctctcctgaa	ttctcagaga
2281	aaacgcctcc	atagtgtctg	ctccaaatta	cacaagttaa	aaagtgaagg	attcatgaaa
2341	gtcctgaagt	gtgaggtgga	gctgatggcc	aggatggcaa	aaaccattga	cagcttcaact
2401	caaaaccaga	caaggctggt	ggttatcatt	gacggcctgg	atgcctgcga	gcaggataaa
2461	gtccttcaga	tgctggacac	tgctccgagt	ctgttttcaa	agggcccttt	tattgccatt
2521	tttgcaagtg	atccgcatat	tatcataaag	gccatcaacc	agaacctgaa	tagtgtgctt
2581	cgtgactcaa	atataaacgg	acatgactac	atgcgcaata	tagttcaact	accagttttt
2641	cttaatagtc	gtgggcttag	caatgcaagg	aaattccttg	taacttcagc	aacaaatggg
2701	gacattacat	gctcagacac	cacaggaaca	caggaggaca	ctgacagaag	agtttcacag
2761	aacagccttg	gggagatgac	aaagcttggt	agcaaaacgg	ctctcaacag	aagggacact
2821	taccgcagaa	gacagatgca	gagaaccatc	acgaggcaga	tgtcctttga	tctcaciaag
2881	ctgctggtca	ctgaggattg	gttcagtga	atcagccctc	agaccatgag	acggctactc
2941	aatattgttt	ctgtgacagg	gcggctacta	agagccaatc	agattacttt	caactgggac
3001	aggctagcca	gctggatcaa	cctcacccag	cagtggccgt	accggacgtc	ttggcttata
3061	ctatatctgg	aagagactga	aggtctccct	gatcagatga	ccttaaagac	catctatgag
3121	agaatatcga	agaatatccc	aacaactaaa	gatgttgagc	cgcttcttga	aattgatgga

Figure 17 (page 2 of 3)

```

3181 gatataagaa actttgaagt cttcttgtct tcaaggactc cagttctcgt ggctcgagat
3241 gtaaagacct ttttgccgtg cacagtaaac ctggatccca aacttcggga gatcatcgca
3301 gatgtccgag ctgcaagaga gcagatcaac ataggaggcc tggcctatcc cccgctccct
3361 ctgcacgaag gccctcctcg gccaccttct gggtagagtc agcctgcata tgtctgctct
3421 tcagcatctt tcaacggggc gttcccaggc ggggttgtgt ctccacagcc ccacagcagt
3481 tactacagcg gcttgtccgg accccagcac cccttctaca acaggccatt ctttgcccca
3541 tacctttaca cgccaaggta ttacctggc ggttcccaac atctcatctc acgttcatca
3601 gtaaaaacga gtttgcccag agatcagaac aatggcctac cgtgtgactc tgggtttaac
3661 aaacagagac aggcagccgt ccctgccaca ggcagctcac tgttactgag ttcaatgacc
3721 gtggatgtcg tatgtgagaa actgagacag atagaagggc tggaccagag catgatgcct
3781 cagtactgca caacaatcaa aaaggcgaac ataaatggcc gggatttgtc tcagtgtaac
3841 attgatgagc tgaagaaaga gatggctatg aactttggag actggcatct cttcagaagc
3901 atggtcttag aaatgagaag tgtggaaagc cagggtggtcc ctgaagacct tcgttttctc
3961 aatgaaaact ccagtgcctc agttcctcat ggagagtctg ctgcgcgctc ttctcacact
4021 gagctgcccc tcacggagct ttccagccag actccataca ccctgaactt cagctttgaa
4081 gaactgaaca cgcttggcct tgatgaggga gcccctagge acagtaacct gagttggcag
4141 tcacaaactc gcagaacccc aagtctctcg agtctcaatt cccaggactc cagtattgaa
4201 atttcaaagc ttactgataa ggtgcaggcc gagtatagag acgcctatag agagtacatt
4261 gctcagatgt cccagttaga agggggcaca gggctgtcca caataagcgg cagatcttct
4321 ccacatagca catactacat aggcagagc tcttcagggg gctccatcca ttctactcta
4381 gaacaggaaa gggggaagga gggtagacta aagcaggagg atgggcggaa gtcattttta
4441 atgaagaggg gggatgtcat agattactcc tcatcagggg tgtccactaa tgaggcctca
4501 cccctggacc ccattactga agaagatgag aaatccgacc agtcaggtag taagctactc
4561 ccaggcaaga aatcctcaga acggcccagt ctcttcagaa cagacttgaa gcttaagggg
4621 ggtggtcttc gctaccagaa actgcccagt gatgaagatg agtcggggac agaagagtca
4681 gataacaccc cactgctcaa agatgataag gacaaaaagg ctgaaggcaa agcagagaga
4741 gtgtgcaagt cccagagca cagtgcagag cccatcagaa cgttcattaa agcaaaagag
4801 tatttgtccg atgccctcct tgacaaaaag gattcctcag attcgggagt gaggtccaat
4861 gagagtcttc ccaaccactc tctccacaat gaggcagcag acgactccca actggaaaag
4921 gcaaacctca tagagcttga agatgagggc cacagtggga agcgggggat gccacacagt
4981 ctgagtggct tgcaagatcc aattatagct cggatgtcca tttgctcgga agacaagaaa
5041 agcccttccg agtgcagcct gattgctagc agccctgaag aaagctggcc tgcgtgccag
5101 aaggcctaca atctaaaccg aacgcccagc actgtgactc tgaacaacaa cactgcaccc
5161 accaacagag ccaatcagaa cttcgatgag atagagggaa tcaggagagc gtcccaggte
5221 attctgaggc ctggcccccag tcccaaccca actgctgtgc agaataaaaa cctgaagagc
5281 atggcgacac agcgaagcca gcgctccagt tacacaaggc tctcgaaaga tgcactctgag
5341 ctacatgcag cctcctcgga gagcacaggc tttggagaag agagagagag cattctctga
5401 tgaggaagga gaagtgaata gagtgtgaca gactttgagt tgaccattcc tcattactca
5461 tcatgtttat gcatcacaga cacctgatag gtggtcaggc actggaagca gacaggaaaag
5521 cctgaaagca cagtcaccag tcattttcat tgtcaccgtc atcccaggca ctttaaagtc
5581 tcactgccac tattacacct ttcttgccac acagctaggc agagcgatga gagcaggctg
5641 aagacaaaat aggtcaaaaag tgtattttaa gttgaaaagg cttaaatcac aaggaaaaaa
5701 acttagagat ctccctttata aactattcat gtaccatttg aggggaagata cagtgttaga
5761 tgaagcagaa accaatttta ttgttagtgt ttaatcttgg tgcagattta taagttttag
5821 agtagccag agactaaaag tgaatactta gcaaattgat agccagtgtt ctatatagga
5881 atcattgctt tttagagggc ttaaagttta agtagaaaat atatactcaa gaaggcgata
5941 aaagctgatg agaaagttag ttagcagaac ccaaagccgt gctgggcccgc ggtgactcat
6001 tagcagagga ggaggggagg gcagtatatt cctgggatac tctctccaga cccagcctgg
6061 cttctgacat catccacctg tgccctcaaa accgtcttag tctgttctgc aactcttaag
6121 tgacatacct aactcagctc atggctaagg aaaaaaaatt aaagttgtcc tggtgattaa
6181 actctggacc tcccacatct aagtcctcag ttgacaaact gcatccccag cccaaacatc
6241 tgttcttcca gagatctgga aagacatggg cataagagtc cagtaaactc ttaatgaagg
6301 tgtcgggaag tgtttccctt aaaggggtta gtgacagtga cactagacaa gtgacagaat

```

## Figure 17 (page 3 of 3)

6361 ttagtttgct gacattaaat cactgccttc tgggcgcgcaa tagtttagta tctgttccta  
6421 aatcagcatt tgagttctgc ggggggttaca aaatggaggc atggtttttaa gaaggaattt  
6481 gaaagcatcc aggaagcaa gaacagagtc ggagctggtg tgaccatgac acttcccatt  
6541 attgttagca aagctaggag aatgaccaca tacatttaag tatcatgtag gctagagcgt  
6601 gcctgtcatc atcatgcatt gagaccgttt gtcctttgat tatagtgtgt ttatgtagac  
6661 cagtcctagc agtatgtcct gagccatggg caggattgta gatggcgagg tggcagggtt  
6721 agtgctcggt gagatgtcca tctctgactg gacctgtgtg gtttacttgt agctgctccg  
6781 tgaaacactg ccttgctggt tgagtgtaat tgagcatttt catgaagtgt atatgttgca  
6841 tgatctttaa caggatattt tagacattaa gaacagcatg atattactga catatgcatt  
6901 agtagtatgc agtatcatgt aatgggtact tcagaggcag tttggtttct ttccagggtg  
6961 gtagttatag ctcacagtca tgggaatcct ttattcacac tgtctgtgaa gtcgatgctg  
7021 gagtgccaaa tgcacacgga gactaggaac aataaagact tgcatgccat tctgactcat  
7081 ggagttttat tcaatagagc aatgataatt ccattctga ttaaaaatta aaacacggaa

Figure 18

1 MSVLISQSVI NYVEEENIPA LKALLEKCKD VDERNECGQT PLMLAAEQGN VEIVKELLKN  
 61 GANCNLEDLD NWTALISASK EGHIIHIVEEL LKSGASLEHR DMGGWTALMW ACYKGRTDVV  
 121 ELLLSHGANG SVTGLYSVYP IIAAAGRGA DIVHLLQNG AKVNCSDKYG TTPLVWAARK  
 181 GHLECVKHLI AMGADVQEG ANSMTALIVA VKGGYTQSVK EILKRNPNVN LTDKDGNTAL  
 241 MIASKEGHIE IVQDLLDAGT YVNIPDRSGD TVLIGAVRGG HVEIVRALLQ KYADIDIRGQ  
 301 DNKTALYWAV EKNATMVRD ILQCNPDTEI CTKDGETPLI KATKMRNIEV VELLIDKGAK  
 361 VSAVDKKGDT PLHVAIRGRS RRLAELLRN PKDGRLLYRP NKAGETPYNI DCSHQKSILT  
 421 QIFGARHLSP TETDGMGLGY DLYSSALADI LSEPTMQPPI CVGLYAQWGS GKSFLKKLE  
 481 DEMKTFAGQO TEPLFQFSWL IVFLTLLLCG GLGLVFAFTV DTNLAIAISL SFLALIYIFF  
 541 IVIYFGGRRE GESWNWAWAL STRLARHIGY LELLFKLMFV NPPELPEQTT KALPVRFLFT  
 601 DYNRLSSVGG ETSLAEMIAT LSDACEREFG FLATRLFRVF RTEESQKKK WKKTCCLPSP  
 661 VIFLFIIVGCI IAGITLLAIF RVDPKHLTVN AILISIASVV GLAFVLCRT WWQVLDLNLN  
 721 SQKRRLHSA SKLHLKSEG FMKVLKCEVE LMARMAKTID SFTQNTRLV VIIDGLDACE  
 781 QDKVLQMLDT VRVLFSGPF IAIASDPHI IIAKAINQNL SVLRDSNING HDYMRNIVHL  
 841 PVFLNSRGLS NARKFLVTS A TNGDITCSDT TGTQEDTDRR VSQNSLGEMT KLGSKTALNR  
 901 RDTYRRRQMQ RTITRQMSFD LTKLLVTEDW FSDISPQTM RLLNIVSVTG RLLRANQITF  
 961 NWDRIASWIN LREQWPYRTS WLILYLEETE GLPDQMTLKT IYERISKNIPT TTKDVEPLLE  
 1021 IDGDIRNFEV FLSSRTPVLV ARDVKTFLPC TVNLDPKLRE IIAADVRAARE QINIGGLAYP  
 1081 PLPLHEGP RP PGYSQPAS VCSSASFNGP FPGGVVSPQ HSSYYSGLSG POHPFFYNRPF  
 1141 FAPYLYTPRY YPGGSQHLIS RSSVKTSLPR QNNGLPCDS GENKQQAAP PATGSSLLLS  
 1201 SMTVDVVECK LRQIEGLDQS MNPQYCTTIK KANINGRVLS QCNIDELKKE MAMNFGDWHL  
 1261 FRSMVLEMRS VESQVVPEDP RFLNENSSAP VPHGESARRS SHTELPLTEL SSQTPYTLNF  
 1321 SFEELNTLGL DEGAPRHSNL SWQSOTRRTP SLSSLNSQDS SIEISKLTDK VQAEYRDAYR  
 1381 EYIAQMSQLE GGTGSSTISG RSSPHSTYI GQSSSGGSIH STLEQERGKE GELKQEDGRK  
 1441 SFLMKRGDVI DYSSSGVSTN EASPLDPITE EDEKSDQSGS KLLPGKKSSE RPSLFQTDLK  
 1501 LKGGGLRYOK LPSEDESGT EESDNTPLLK DDKDKKAEGK AERVCKSPEH SAEPITRFIK  
 1561 AKEYLSDALL DKKDSSDSGV RSNESSPNHS LHNEAADDSQ LEKANLIELE DEGHSGKRG  
 1621 PHLSGLQDP IIAARMSICSE DKKSPSECSL IASSPEESWP ACQKAYNLNR TPSTVTLNN  
 1681 TAPTNRANQN FDEIEGIRET SQVILRPGPS PNPTAVQEN LKSMHAKRSQ RSSYTRLSKD  
 1741 ASELHAASSE STGFGEERES IL



Figure 19

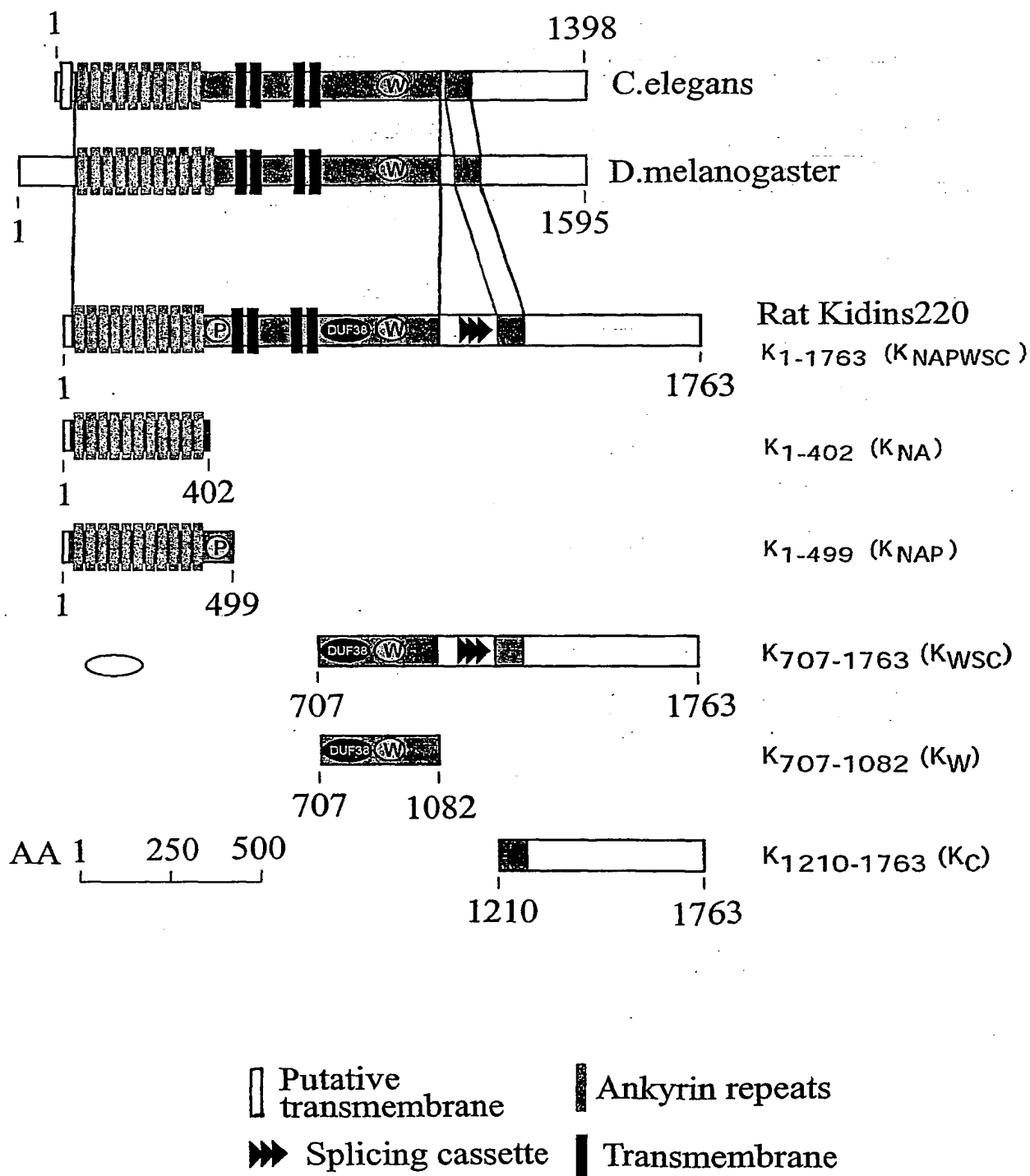


Figure 20

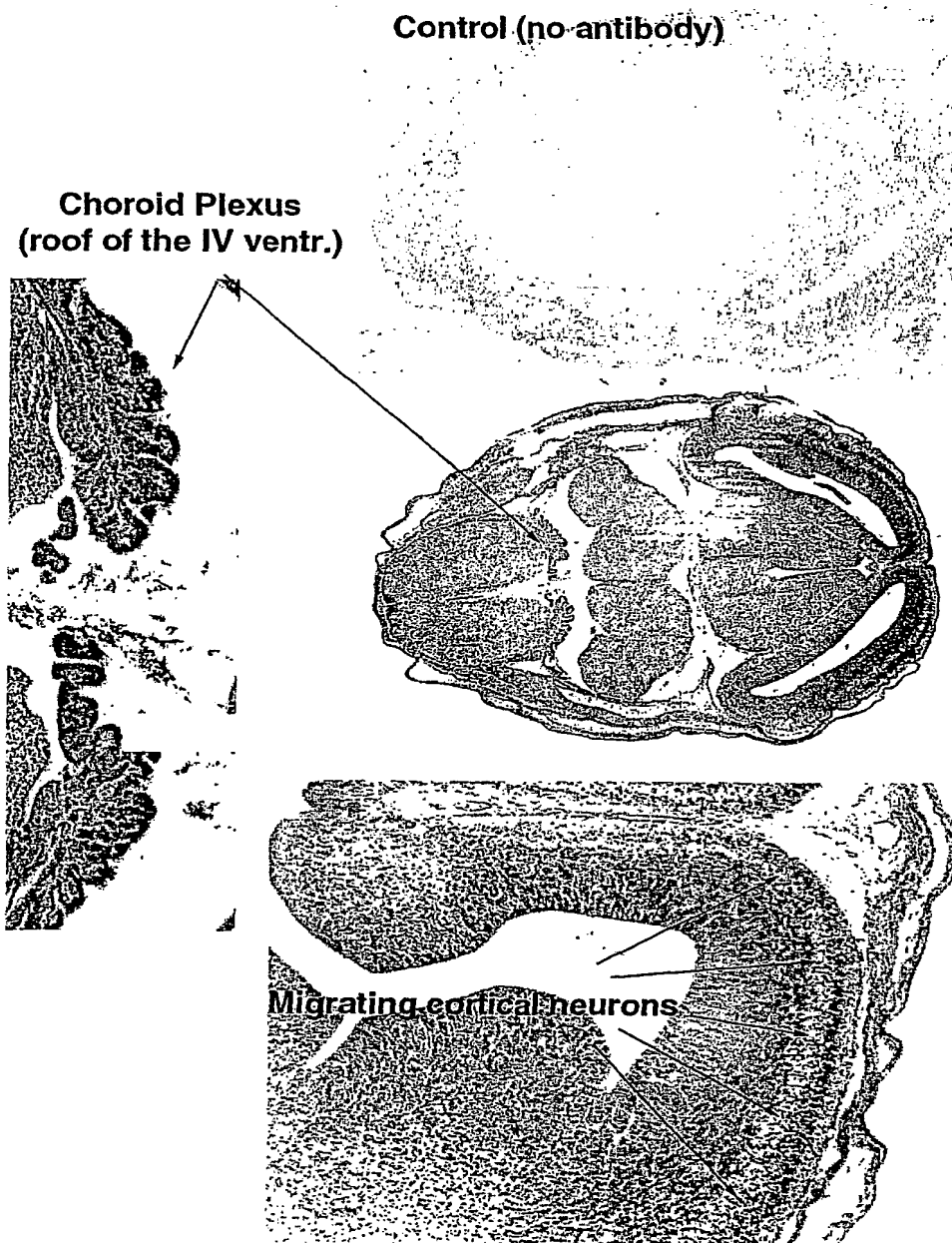


Figure 21

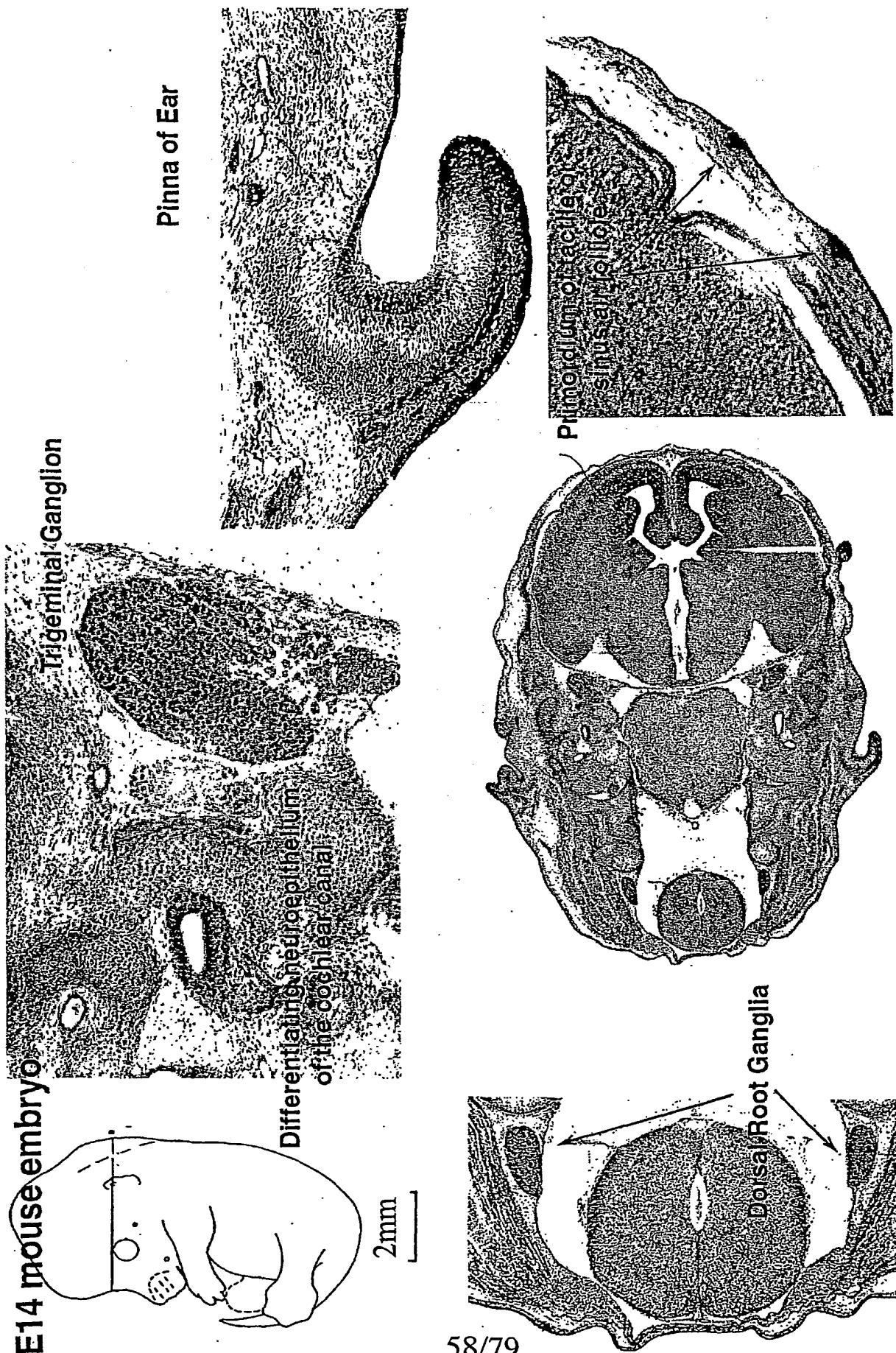


Figure 22 (page 1 of 4)



Serial Analysis of Gene Expression  
Tag to Gene Mapping



CGAP

UniGene

OMIM

PubMed

Entrez

BLAST

## SAGE Gene to Tag Mapping

UniGene cluster id:

**Hs.9873 : likely homolog of rat kinase D-interacting substance of 220 kDa; KIAA1250 protein**

SAGE library data and reliable tag summary:

Reliable tags found in SAGE libraries:

CTAAAAAAAA

<i>Library name</i>	<i>Tags per million</i>	<i>Tag counts</i>	<i>Total tags</i>
<u>SAGE Duke 1273</u>	51	2	38836
<u>SAGE Duke thalamus</u>	246	6	24371
<u>SAGE Duke H1020</u>	114	6	52371
<u>SAGE Duke BB542 normal cerebellum</u>	293	12	40909
<u>SAGE 293-CTRL</u>	46	2	43442
<u>SAGE HCT116</u>	165	10	60322
<u>SAGE Caco 2</u>	487	30	61601
<u>SAGE Chen LNCaP</u>	176	11	62267
<u>SAGE Chen LNCaP no-DHT</u>	61	4	64631
<u>SAGE Chen Normal Pr</u>	105	7	66193
<u>SAGE Chen Tumor Pr</u>	204	14	68384
<u>SAGE CAPAN1</u>	26	1	37926
<u>SAGE CAPAN2</u>	129	3	23222
<u>SAGE HS766T</u>	95	1	10467
<u>SAGE Panc1</u>	40	1	24879
<u>SAGE H126</u>	30	1	32420
<u>SAGE Duke H54 lacZ</u>	327	22	67101
<u>SAGE Duke H54 EGFRvIII</u>	454	26	57164
<u>SAGE Duke H392</u>	139	8	57529
<u>SAGE Duke GBM H1110</u>	870	61	70061
<u>SAGE SW837</u>	409	25	60986
<u>SAGE RKO</u>	249	13	52064
<u>SAGE CPDR LNCaP-C</u>	240	10	41590
<u>SAGE CPDR LNCaP-T</u>	339	15	44122
<u>SAGE 293-IND</u>	163	4	24481

59/79

SUBSTITUTE SHEET (RULE 26)

Figure 22 (page 2 of 4)

<u>SAGE PR317 normal prostate</u>	858	51	59419
<u>SAGE PR317 prostate tumor</u>	491	32	65109
<u>SAGE pooled GBM</u>	355	22	61841
<u>SAGE BB542 whitematter</u>	379	36	94806
<u>SAGE NHA(5th)</u>	344	18	52196
<u>SAGE normal pool(6th)</u>	174	11	63064
<u>SAGE NC1</u>	39	2	50115
<u>SAGE NC2</u>	60	3	49552
<u>SAGE Panc 91-16113</u>	265	9	33941
<u>SAGE Panc 96-6252</u>	335	12	35745
<u>SAGE OV1063-3</u>	179	7	38938
<u>SAGE Tu102</u>	121	7	57636
<u>SAGE Tu98</u>	81	4	49005
<u>SAGE Duke mhh-1</u>	164	8	48488
<u>SAGE SciencePark MCF7 control 3h</u>	169	1	5903
<u>SAGE SciencePark MCF7 Control 0h</u>	278	17	61079
<u>SAGE SciencePark MCF7 estradiol 3h</u>	150	9	59978
<u>SAGE SciencePark MCF7 estradiol 10h</u>	82	5	60435
<u>SAGE lacZ</u>	161	3	18528
<u>SAGE 95-347</u>	89	6	67240
<u>SAGE 95-259</u>	152	6	39473
<u>SAGE 95-260</u>	132	6	45179
<u>SAGE 95-348</u>	429	26	60484
<u>SAGE Ped GBM1062</u>	1868	112	59935
<u>SAGE Duke H341</u>	112	5	44563
<u>SAGE HOSE 4</u>	433	21	48413
<u>SAGE ES2-1</u>	31	1	31502
<u>SAGE OVP-5</u>	462	5	10802
<u>SAGE Medullo 3871</u>	392	17	43274
<u>SAGE PrCA-1</u>	1427	13	9105
<u>SAGE normal prostate</u>	684	9	13148
<u>SAGE LN-1</u>	525	12	22835
<u>SAGE PERITO-13</u>	446	24	53728
<u>SAGE MouseP8 PGCP</u>	114	7	61240
<u>SAGE Meso-12</u>	685	24	35032
<u>SAGE OC14</u>	400	7	17476

60/79

SUBSTITUTE SHEET (RULE 26)

Figure 22 (page 3 of 4)

<u>SAGE LNCaP</u>	353	8	22637
<u>SAGE OVT-6</u>	188	8	42336
<u>SAGE H1126</u>	514	9	17501
<u>SAGE OVT-7</u>	382	21	54914
<u>SAGE HMEC-B41</u>	709	1	1410
<u>SAGE MDA453</u>	211	4	18924
<u>SAGE SKBR3</u>	735	6	8153
<u>SAGE Duke HMVEC</u>	228	12	52532
<u>SAGE Duke HMVEC+VEGF</u>	103	6	57928
<u>SAGE mammary epithelium</u>	467	23	49167
<u>SAGE DCIS</u>	242	10	41230
<u>SAGE Duke 757</u>	358	7	19503
<u>SAGE normal cerebellum</u>	254	13	51135
<u>SAGE OVT-8</u>	655	22	33575
<u>SAGE Duke 40N</u>	140	1	7142
<u>SAGE Duke 48N</u>	165	2	12091
<u>SAGE ML10-10</u>	228	13	56943
<u>SAGE Duke H247 normal</u>	363	22	60543
<u>SAGE Duke H247 Hypoxia</u>	347	25	71937
<u>SAGE Duke-H988</u>	285	8	28015
<u>SAGE Duke post crisis fibroblasts</u>	180	4	22207
<u>SAGE Duke precrisis fibroblasts</u>	113	1	8797
<u>SAGE DCIS 2</u>	242	7	28888
<u>SAGE Br N</u>	213	8	37558
<u>SAGE A+</u>	130	4	30551
<u>SAGE IOSE29-11</u>	226	11	48498
<u>SAGE Duke H1043</u>	195	15	76673

CTGCTTCCT

<i>Library name</i>	<i>Tags per million</i>	<i>Tag counts</i>	<i>Total tags</i>
<u>SAGE Caco 2</u>	16	1	61601
<u>SAGE CAPAN1</u>	26	1	37926
<u>SAGE BB542 whitematter</u>	21	2	94806
<u>SAGE NC2</u>	20	1	49552
<u>SAGE 95-348</u>	16	1	60484
<u>SAGE HOSE 4</u>	20	1	48413
<u>SAGE PERITO-13</u>	18	1	53728
<u>SAGE OC14</u>	57	1	17476
<u>SAGE Duke HMVEC+VEGF</u>	17	1	57928




Figure 22 (page 4 of 4)

<u>GCTCACCAGC</u>	<u>SAGE mammary epithelium</u>	20	1	49167
	<u>SAGE normal cerebellum</u>	19	1	51135
	<u>SAGE DCIS 2</u>	34	1	28888
	<i>Library name</i>	<i>Tags per million</i>	<i>Tag counts</i>	<i>Total tags</i>
	<u>SAGE Caco 2</u>	16	1	61601
	<u>SAGE Chen Tumor Pr</u>	14	1	68384
	<u>SAGE Duke H54 EGFRvIII</u>	17	1	57164
	<u>SAGE Duke H392</u>	17	1	57529
	<u>SAGE pooled GBM</u>	16	1	61841
	<u>SAGE normal pool(6th)</u>	31	2	63064
	<u>SAGE Duke mhh-1</u>	20	1	48488
	<u>SAGE OVT-7</u>	18	1	54914
	<u>SAGE Duke H247 normal</u>	16	1	60543
	<u>SAGE Duke H247 Hypoxia</u>	27	2	71937
	<u>SAGE Duke H1043</u>	26	2	76673

Number of SAGE libraries: 95

Total tags in all SAGE libraries: 3888724

Summary of tags found in this cluster:

<i>Tag</i>	<i>Cid- &gt;tag frequency</i>	<i>Get seqs</i>
cDNA: well-characterized		
<u>CTAAAAAAAA</u>	1/4	
<u>CTGCTTCCCT</u>	1/4	
<u>GCTCACCAGC</u>	2/4	

NLM

| NIH

| SAGEnet

| GGEG

| NCBI Help

| Disclaimer

**Figure 23**

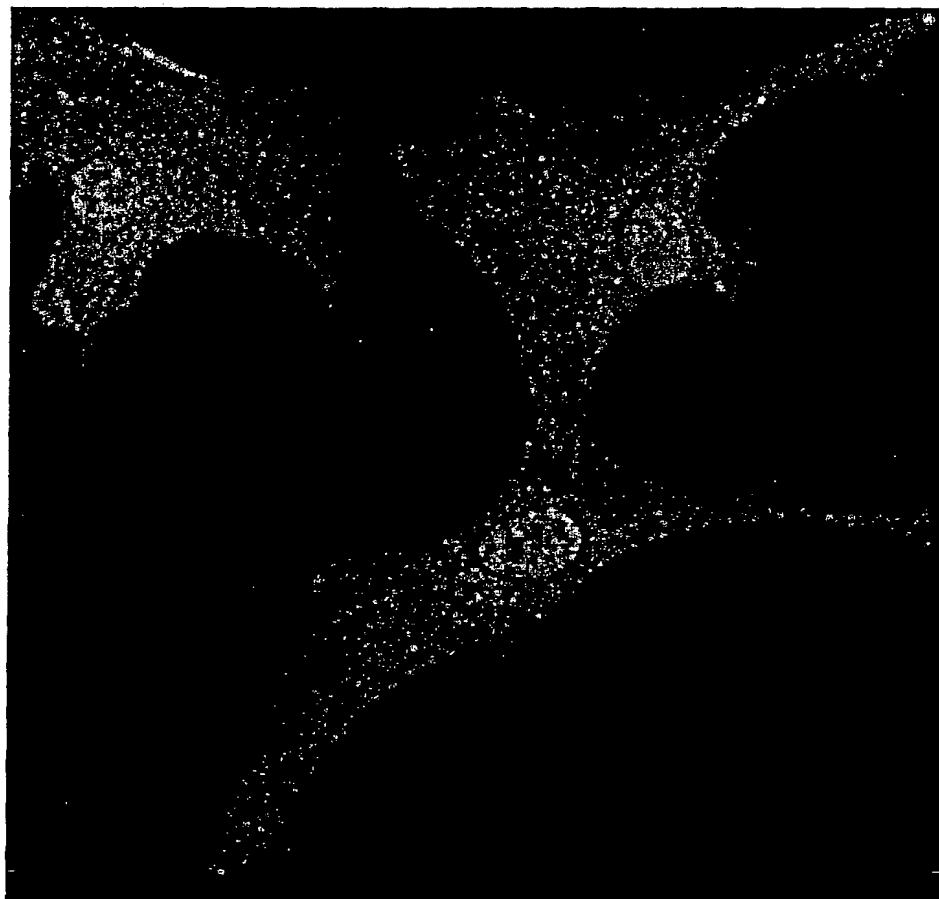




Figure 24 (page 1 of 12)

1 gtcctctgacaaggtctgcggccctgtgagctggctctacagagccgagg 50  
|||||  
1 gtcctctgacaaggtctgcggccctgtgagctggctctacagagccgagg 50  
51 gcgagtcctcagccaggccgagatacgggtgaattaaagtcggcaataga 100  
|||||  
51 gcgagtcctcagccaggccgagatacgggtgaattaaagtcggcaataga 100  
101 tttctgtgaaaATGTCAGTTCTTATATCACAGAGTGTGATAAATTATGTG 150  
|||||  
101 tttctgtgaaaatgtcagttcttatatcacagagtgtgataaattatgtg 150  
151 GAAGAAGAGAATATTCCTGCTCTGAAAGCTCTTCTTGAGAAGTGCAAAGA 200  
|||||  
151 gaagaagagaatattcctgctctgaaagctcttcttgagaagtgcaaaga 200  
201 TGTAGATGAGAGGAACGAGTGTGGCCAGACGCCGCTGATGCTGGCTGCGG 250  
|||||  
201 tgtagatgagaggaacgagtggtggccagacgccgctgatgctggctgcgg 250  
251 AGCAGGGCAATGTGGAATAAGTGAAAGAACTGTTAAAGAATGGAGCCAAC 300  
|||||  
251 agcagggcaatgtggaaatagtgaagaactgttaaagaatggagccaac 300  
301 TGCAATCTGGAAGATCTGGATAACTGGACTGCCCTTATATCGGCATCTAA 350  
|||||  
301 tgcaatctggaagatctggataactggactgcccttatatcggcatctaa 350  
351 GGAGGGGCACATCCACATTGTGGAGGAGCTGCTGAAGAGCGGGGCCAGTC 400  
|||||  
351 ggaggggcacatccacattgtggaggagctgctgaagagcggggccagtc 400  
401 TGGAGCACCGAGACATGGGTGGATGGACAGCTCTCATGTGGGCGTGCTAC 450  
|||||  
401 tggagcaccgagacatgggtggatggacagctctcatgtgggcgtgctac 450  
451 AAAGGCAGGACCGACGTGGTCGAGCTGCTTCTTCTCATGGTGCCAACCC 500  
|||||  
451 aaaggcaggaccgacgtggtcgagctgcttcttctcatggtgccaaccc 500  
501 AAGTGTCACTGGTTTGCAGTATAGTGTCTACCCAATCATCTGGGCAGCGG 550  
|||||  
501 aagtgtcactggttt...gtatagtgctacccaatcatctgggcagcgg 547

Figure 24 (page 2 of 12)

551 GCAGAGGCCATGCCGACATAGTGCATCTTCTACTGCAGAATGGTGCGAAA 600  
|||||  
548 gcagaggccatgccgacatagtgcacatcttctactgcagaatgggtgcgaaa 597  
601 GTCAATTGCTCCGACAAGTACGGAACAACCCCTTGGTATGGGCTGCCCG 650  
|||||  
598 gtcaattgctccgacaagtacggaacaaccccttgggtatgggctgcccg 647  
651 AAAGGGTCATTTGGAATGTGTGAAGCACTTGTGGCCATGGGAGCTGATG 700  
|||||  
648 aaagggtcatttggaatgtgtgaagcacttgttggccatgggagctgatg 697  
701 TTGATCAAGAAGGAGCCAATTTCGACGACTGCACTTATCGTGGCAGTGAAA 750  
|||||  
698 ttgatcaagaaggagccaattcgatgactgcacttatcgtggcagtgaaa 747  
751 GGAGGCTACACACAGTCCGTGAAAGAGATTTTGAAAAGGAATCCAAATGT 800  
|||||  
748 ggaggctacacacagtcctgtgaaagagattttgaaaaggaatccaaatgt 797  
801 GAACCTAAGTACAGGATGGAAACACAGCTCTGATGATCGCATCGAAGG 850  
|||||  
798 gaacctaactgacaaggatggaaacacagctctgatgatcgcatcgaagg 847  
851 AGGGACATATTGAGATTGTGCAGGACCTGCTGGACGCCGGAACATACGTG 900  
|||||  
848 agggacatattgagattgtgcaggacctgctggacgccggaacatacgtg 897  
901 AACATACCTGACAGGAGTGGGGATACTGTATTGATCGGCGCTGTGAGAGG 950  
|||||  
898 aacatacctgacaggagtggggatactgtattgatcggcgctgtgagagg 947  
951 TGGTCATGTTGAAATTGTTTCGAGCACTTCTCCAAAAATATGCTGATATAG 1000  
|||||  
948 tggtcattgttgaaattgttcgagcacttctccaaaaatattgctgatatag 997  
1001 ACATTAGAGGGCAGGACAATAAGACCGCGCTGTATTGGGCCGTTGAAAAG 1050  
|||||  
998 acattagagggcaggacaataagaccgcgctgtattgggccgttgaaaag 1047  
1051 GGAAATGCGACAATGGTGAGGGATATCCTGCAGTGCAATCCGGACACGGA 1100  
|||||  
1048 ggaaatgcgacaatgggtgagggatatcctgcagtgcaatccggacacgga 1097  
1101 AATATGCACGAAGGATGGTGAGACACCACTGATAAAAGCCACCAAGATGA 1150  
|||||  
1098 aatatgcacgaaggatgggtgagacaccactgataaaagccaccaagatga 1147

1151	GAAATACTGAAGTGGTGGAGCTGCTGCTGGATAAGGGGGGCTAAGGTGTCT	1200
1148	gaaatattgaagtggtagctgctgctggataaggggggctaaggtgtct	1197
1201	GCTGTGGACAAGAAAGGGGACACTCCCCTGCACGTTGCCATCCGAGGGAG	1250
1198	gctgtggacaagaaaggggacactcccctgcacgttgccatccgagggag	1247
1251	GAGCCGGAGACTGGCCGAACCTTCTTTTACGGAACCCCAAAGATGGACGGT	1300
1248	gagccggagactggccgaacttcttttacggaaccccaaagatggacggt	1297
1301	TACTCTACAGACCCAACAAAGCAGGCGAGACTCCCTACAACATTGACTGC	1350
1298	tactctacagacccaacaaagcaggcgagactccctacaacattgactgc	1347
1351	AGCCACCAGAAAAGCATTTTAACTCAAATATTTGGAGCCAGACACTTGTC	1400
1348	agccaccagaaaagcattttaactcaaataatttgagccagacacttgtc	1397
1401	TCCTACGGAAACAGATGGCGACATGCTTGGTTATGATTGTATAGCAGTG	1450
1398	tcctacggaaacagatggcgacatgcttggttatgatttgtatagcagtg	1447
1451	CCCTGGCCGATATTCTCAGTGAACCCACCATGCAACCACCCATTTGTGTG	1500
1448	ccctggccgatattctcagtgaaccacccatgcaaccacccatttgtgtg	1497
1501	GGCTTGTATGCACAGTGGGGAAGTGGGAAGTCTTTCTTACTCAAGAACT	1550
1498	ggcttgtatgcacagtggggaagtgggaagtctttcttactcaagaaact	1547
1551	AGAAGATGAGATGAAGACGTTTGCAGGACAGCAGACCGAGCCCCCTCTTCC	1600
1548	agaagatgagatgaagacgtttgcaggacagcagaccgagcccctcttcc	1597
1601	AGTTCTCGTGGCTCATAGTGTTCCCTCACCCCTGCTGCTGTGTGGCGGGCTC	1650
1598	agttctcgtggctcatagtgttcctcacccctgctgctgtgtggcgggctc	1647
1651	GGCCTAGTGTTTGCTTTCACGGTTGATACAAACCTCGCCATAGCCATTTC	1700
1648	ggcctagtgtttgctttcacggttgatacaaacctcgccatagccatttc	1697
1701	ATTGAGTTTCCTGGCTCTCATATATATATTCTTCATCGTCATTTACTTCG	1750
1698	attgagtttcctggctctcatatatatatattcttcacgctcatttacttcg	1747

Figure 24 (page 4 of 12)

1751 GTGGACGGCGGGAAGGAGAGAGTTGGAACCTGGGCCTGGGCTCTCAGTACC 1800  
|||||  
1748 gtggacggcggaaggagagagttggaactgggcctgggctctcagtacc 1797  
1801 AGACTGGCAAGACACATTGGGTACCTGGAGCTCCTCTTCAAACCTGATGTT 1850  
|||||  
1798 agactggcaagacacattgggtacctggagctcctcttcaaactgatggt 1847  
1851 TGTGAACCCACCTGAACCTGCGGAGCAAACCTACAAAGGCTTTACCTGTGA 1900  
|||||  
1848 tgtgaaccacactgaactgccggagcaaactacaaaggctttacctgtga 1897  
1901 GGTTTTTGTTTACAGATTACAATAGACTCTCCAGTGTGGGAGGAGAACT 1950  
|||||  
1898 ggtttttgtttacagattacaatagactctccagtgtgggaggagaaact 1947  
1951 TCCTTGGCTGAAATGATCGCGACCCTCTCAGATGCTTGTGAGAGAGAATT 2000  
|||||  
1948 tccttggctgaaatgatcgcgaccctctcagatgcttgtgagagagaatt 1997  
2001 CGGCTTTTTGGCTACCCGTCTTTTTCGAGTCTCAGGACTGAAGAGTCTC 2050  
|||||  
1998 cggctttttggctacccgtcttttttcgagtcttcaggactgaagagtctc 2047  
2051 AGGGTAAAAAGAAATGGAAAAGAACATGCTGCCTCCCCTCTTTTGTCTATC 2100  
|||||  
2048 agggtaaaaagaaatggaaaaaacatgctgcctcccctcttttgtcatc 2097  
2101 TTTCTTTTTATTGTTGGCTGCATTATTGCTGGGATCACCTTCTGGCTAT 2150  
|||||  
2098 tttctttttattgttggctgcattattgctgggatcaccttctggctat 2147  
2151 ATTCAGAGTTGACCCGAAACATCTGACAGTGAATGCTATCCTCATTTCTA 2200  
|||||  
2148 attcagagttgaccgaaacatctgacagtgaatgctatcctcatttcta 2197  
2201 TCGCATCTGTGCTGGGGTTGGCCTTCGTGCTCAATTGTCGAACATGGTGG 2250  
|||||  
2198 tcgcatctgtcgtggggttggccttcgtgctcaattgtcgaacatggtgg 2247  
2251 CAAGTGCTGGACTCTCTCCTGAATTCTCAGAGAAAACGCCTCCATAGTGC 2300  
|||||  
2248 caagtgctggactctctcctgaattctcagagaaaacgcctccatagtgc 2297  
2301 TGCCTCCAAATTACACAAGTTAAAAAGTGAAGGATTCATGAAAGTCCTGA 2350  
|||||  
2298 tgcctccaaattacacaagttaaaaagtgaaggattcatgaaagtcctga 2347

Figure 24 (page 5 of 12)

2351 AGTGTGAGGTGGAGCTGATGGCCAGGATGGCAAAAACCATTTGACAGCTTC 2400  
|||||  
2348 agtgtgaggtggagctgatggccaggatggcaaaaaccattgacagcttc 2397  
2401 ACTCAAAAACCAGACAAGGCTGGTGGTTATCATTGACGGCCTGGATGCCTG 2450  
|||||  
2398 actcaaaaccagacaaggctggtggttatcattgacggcctggatgcctg 2447  
2451 CGAGCAGGATAAAGTCCTTCAGATGCTGGACACTGTCCGAGTTCTGTTTT 2500  
|||||  
2448 cgagcaggataaagtccttcagatgctggacactgtccgagttctgtttt 2497  
2501 CAAAGGGCCCTTTTATTGCCATTTTTGCAAGTGATCCGCATATTATCATA 2550  
|||||  
2498 caaagggccctttttattgccatTTTTgcaagtgatccgcatattatcata 2547  
2551 AAGGCCATCAACCAGAACCTGAATAGTGTGCTTCGTGACTCAAATATAAA 2600  
|||||  
2548 aagggcatcaaccagaacctgaatagtgtgcttcgtgactcaaataaaa 2597  
2601 CGGACATGACTACATGCGCAATATAGTTCACCTACCAGTTTTTCTTAATA 2650  
|||||  
2598 cggacatgactacatgcgcaatatagttcacttaccagtttttcttaata 2647  
2651 GTCGTGGGCTTAGCAATGCAAGGAAATTCCTTGTAACCTCAGCAACAAAT 2700  
|||||  
2648 gtcgtgggcttagcaatgcaaggaaattccttgtaacttcagcaacaaat 2697  
2701 GGGGACATTACATGCTCAGACACCACAGGAACACAGGAGGACACTGACAG 2750  
|||||  
2698 ggggacattacatgctcagacaccacaggaacacaggaggacactgacag 2747  
2751 AAGAGTTTCACAGAACAGCCTTGGGGAGATGACAAAGCTTGGGAGCAAAA 2800  
|||||  
2748 aagagtttcacagaacagccttggggagatgacaaagcttgggagcaaaa 2797  
2801 CGGCTCTCAACAGAAGGGACACTTACCGCAGAAGACAGATGCAGAGAACC 2850  
|||||  
2798 cggctctcaacagaaggggacacttaccgcagaagacagatgcagagaacc 2847  
2851 ATCACGAGGCAGATGTCCTTTGATCTCACAAAGCTGCTGGTCACTGAGGA 2900  
|||||  
2848 atcacgaggcagatgtcctttgatctcacaagctgctggtcactgagga 2897  
2901 TTGGTTCAGTGACATCAGCCCTCAGACCATGAGACGGCTACTCAATATTG 2950  
|||||  
2898 ttggttcagtgacatcagccctcagaccatgagacggctactcaatattg 2947

## Figure 24 (page 6 of 12)

2951 TTTCTGTGACAGGGCGGCTACTAAGAGCCAATCAGATTACTTTCAACTGG 3000  
|||||  
2948 tttctgtgacagggcggctactaagagccaatcagattactttcaactgg 2997  
3001 GACAGGCTAGCCAGCTGGATCAACCTCACCGAGCAGTGGCCGTACCGGAC 3050  
|||||  
2998 gacaggctagccagctggatcaacctcaccgagcagtggccgtaccggac 3047  
3051 GTCTTGGCTTATACTATATCTGGAAGAGACTGAAGGTCTCCCTGATCAGA 3100  
|||||  
3048 gtcttggcttatactatatactggaagagactgaaggtctccctgatcaga 3097  
3101 TGACCTTAAAGACCATCTATGAGAGAATATCGAAGAATATCCCAACAAC 3150  
|||||  
3098 tgaccttaaagaccatctatgagagaatatcgaagaatatcccaacaact 3147  
3151 AAAGATGTTGAGCCGCTTCTTGAAATTGATGGAGATATAAGAACTTTGA 3200  
|||||  
3148 aaagatgttgagccgcttcttgaaattgatggagatataagaaactttga 3197  
3201 AGTCTTCTTGTCTTCAAGGACTCCAGTTCTCGTGGCTCGAGATGTAAAGA 3250  
|||||  
3198 agtcttcttgtcttcaaggactccagttctcgtggctcgagatgtaaaga 3247  
3251 CCTTTTTGCCGTGCACAGTAAACCTGGATCCCAAACCTTCGGGAGACCATC 3300  
|||||  
3248 cctttttgccgtgcacagtaaacctggatcccaaacttcgggagatcatc 3297  
3301 GCAGATGTCCGAGCTGCAAGAGAGCAGATCAACATAGGAGGCCTGGCCTA 3350  
|||||  
3298 gcagatgtccgagctgcaagagagcagatcaacataggaggcctggccta 3347  
3351 TCCCCCGCTCCCTCTGCACGAAGGCCCTCCTCGGCCACCTTCTGGGTACA 3400  
|||||  
3348 tcccccgctccctctgcacgaaggccctcctcggccaccttctgggtaca 3397  
3401 GTCAGCCTGCATCTGTCTGCTCTTCAGCATCTTTCAACGGGCCGTTCCCA 3450  
|||||  
3398 gtcagcctgcatctgtctgctcttcagcatctttcaacgggccggttcca 3447  
3451 GGCGGGGTTGTGTCTCCACAGCCCCACAGCAGTTACTACAGCGGCTTGTC 3500  
|||||  
3448 ggcgggggttgtgtctccacagccccacagcagttactacagcggttgtc 3497  
3501 CGGACCCCAGCACCCCTTCTACAACAGGCCATTCTTTGCCCCATACCTTT 3550  
|||||  
3498 cggaccccagcacccttctacaacaggccattctttgccccataccttt 3547

## Figure 24 (page 7 of 12)

3551 ACACGCCAAGGTATTACCCTGGCGGTTCCCAACATCTCATCTCACGTTCA 3600  
|||||  
3548 acacgccaaggtattaccctggcggttcccaacatctcatctcacgttca 3597  
3601 TCAGTAAAAGCGAGTTTGCCCGAGAGATCAGAACAAATGGCCTACCGTGTGA 3650  
|||||  
3598 tcagtataaacgagtttgcccagagatcagaacaatggcctaccgtgtga 3647  
3651 CTCTGGGTTTAACAAACAGAGACAGGGCAGCCGTCCCTGCCACAGGCAGC 3700  
|||||  
3648 ctctgggtttaacaaacagagaca.ggcagccgtccctgccacaggcagc 3696  
3701 TCACTGTTACTGAGTTCAATGACCGTGGATGTCGT.TGTGAGAACTGAG 3749  
|||||  
3697 tcactgttactgagttcaatgaccgtggatgtcgtatgtgagaaactgag 3746  
3750 ACAGATAGAAGGGCTGGACCAGAGCATGATGCCTCAGTACTGCACAACAA 3799  
|||||  
3747 acagatagaagggctggaccagagcatgatgcctcagtactgcacaacaa 3796  
3800 TCAAAAAGGCGAACATAAATGGCCGGGTATTGTCTCAGTGTAACATTGAT 3849  
|||||  
3797 tcaaaaaggcgaacataaatggccgggtattgtctcagtgtaacattgat 3846  
3850 GAGCTGAAGAAAGAGATGGCTATGAACTTTGGAGACTGGCATCTCTTCAG 3899  
|||||  
3847 gagctgaagaaagagatggctatgaactttggagactggcatctcttcag 3896  
3900 AAGCATGGTCTTAGAAATGAGAAGTGTGGAAAGCCAGGTGGTCCCTGAAG 3949  
|||||  
3897 aagcatggtcttagaaatgagaagtgtggaaagccaggtggtccctgaag 3946  
3950 ACCCTCGTTTTCTCAATGAAAACCTCCAGTGCCCCAGTTCCTCATGGAGAG 3999  
|||||  
3947 accctcgttttctcaatgaaaactccagtgccccagttcctcatggagag 3996  
4000 TCTGCTCGCCGCCCTTCTCACACTGAGCTGCCCCTCACGGAGCTTTCCAG 4049  
|||||  
3997 tctgctcgccgctcttctcacactgagctgcccctcacggagctttccag 4046  
4050 CCAGACTCCATACACCCTGAACTTCAGCTTTGAAGAACTGAACACGCTTG 4099  
|||||  
4047 ccagactccatacacccctgaacttcagctttgaagaactgaacacgcttg 4096  
4100 GCCTTGATGAGGGTGCCCCCTCGTCACAGTAATCTAAGTTGGCAGTCACAA 4149  
|||||  
4097 gccttgatgagggtagccctaggcacagtaacctgagttggcagtcacaa 4146

Figure 24 (page 8 of 12)

4150 ACTCGCAGAACCCCAAGTCTCTCGAGTCTCAATTCCCAGGACTCCAGTAT 4199  
|||||  
4147 actcgcagaaccccaagtctctcgagtctcaattcccaggactccagtat 4196  
4200 TGAAATTTCAAAGCTTACTGATAAGGTGCAGGCCGAGTATAGAGACGCCT 4249  
|||||  
4197 tgaaatttcaaagcttactgataaggtgcaggccgagtatagagacgcct 4246  
4250 ATAGAGAGTACATTGCTCAGATGTCCCAGTTAGAAGGGGGCACAGGGTCG 4299  
|||||  
4247 atagagagtacattgctcagatgtcccagttagaagggggcacaggggtcg 4296  
4300 TCCACAATAAGCGGCAGATCTTCTCCACATAGCACATACTACATAGGCCA 4349  
|||||  
4297 tccacaataagcggcagatcttctccacatagcacatactacataggcca 4346  
4350 GAGCTCTTCAGGGGGCTCCATCCATTCTACTCTAGAACAGGAAAGGGGGA 4399  
|||||  
4347 gagctcttcaggggggtccatccattctactctagaacaggaaaggggga 4396  
4400 AGGAGGGGTGAGCTAAAGCAGGAGGATGGGCGGAAGTCATTTTAAATGAAG 4449  
|||||  
4397 aggaggggtgagctaaagcaggaggatgggcggaagtcatttttaatgaag 4446  
4450 AGGGGGGATGTCATAGATTACTCCTCATCAGGGGTGTCCACTAATGAGGC 4499  
|||||  
4447 aggggggatgtcatagattactcctcatcaggggtgtccactaatgaggc 4496  
4500 CTCACCCCTGGACCCCATTAAGTGAAGAAGATGAGAAATCCGACCAGTCAG 4549  
|||||  
4497 ctcacccctggaccccatctactgaagaagatgagaaatccgaccagtcag 4546  
4550 GTAGTAAGCTACTCCCAGGCAAGAAATCCTCAGAACGGCCCAGTCTCTTC 4599  
|||||  
4547 gtagtaagctactcccaggcaagaaatcctcagaacggcccagtcctcttc 4596  
4600 CAGACAGACTTGAAGCTTAAGGGGGGTGGTCTTCGCTACCAGAACTGCC 4649  
|||||  
4597 cagacagacttgaagcttaaggggggtggtcttcgctaccagaaactgcc 4646  
4650 CAGTGATGAAGATGAGTCGGGGACAGAAGAGTCAGATAACACCCCACTGC 4699  
|||||  
4647 cagtgatgaagatgagtcggggacagaagagtcagataacaccccactgc 4696  
4700 TCAAAGATGATAAGGACAAAAGGCTGAAGGCAAAGCAGAGAGAGTGTGC 4749  
|||||  
4697 tcaaagatgataaggacaaaaggctgaaggcaaagcagagagagtggtgc 4746



Figure 24 (page 9 of 12)

4750 AAGTCCCCAGAGCACAGTGCCGAGCCCATCAGAACGTTTCATTAAAGCAAA 4799  
|||||  
4747 aagtccccagagcacagtgccgagcccattcagaacgttcattaaagcaaa 4796  
4800 AGAGTATTTGTCCGATGCCCTCCTTGACAAAAAGGATTCCTCAGATTTCGG 4849  
|||||  
4797 agagtatttgtccgatgccctccttgacaaaaaggattcctcagattcgg 4846  
4850 GAGTGAGGTCCAATGAGAGTTCTCCCAACCACTCTCTCCACAATGAGGCA 4899  
|||||  
4847 gagtgaggtccaatgagagttctcccaaccactctctccacaatgaggca 4896  
4900 GCAGACGACTCCCAACTGGAAAAGGCAAACTCATAGAGCTTGAAGATGA 4949  
|||||  
4897 gcagacgactcccaactggaaaaggcaaacctcatagagcttgaagatga 4946  
4950 GGGCCACAGTGGGAAGCGGGGGATGCCACACAGTCTGAGTGGCTTGCAAG 4999  
|||||  
4947 gggccacagtgggaagcgggggatgccacacagtctgagtggcttgcaag 4996  
5000 ATCCAATTATAGCTCGGATGTCCATTTGCTCGGAAGACAAGAAAAGCCCT 5049  
|||||  
4997 atccaattatagctcggatgtccatttgctcgggaagacaagaaaagccct 5046  
5050 TCCGAGTGCAGCCTGATTGCTAGCAGCCCTGAAGAAAGCTGGCCTGCGTG 5099  
|||||  
5047 tccgagtgcagcctgattgctagcagccctgaagaaagctggcctgctg 5096  
5100 CCAGAAGGCCTACAATCTAAACCGAACGCCCAGCACTGTGACTCTGAACA 5149  
|||||  
5097 ccagaaggcctacaatctaaaccgaacgccagcactgtgactctgaaca 5146  
5150 ACAACACTGCACCCACCAACAGAGCCAATCAGAACTTCGATGAGATAGAG 5199  
|||||  
5147 acaacactgcacccaccaacagagccaatcagaacttcgatgagatagag 5196  
5200 GGAATCAGGGAGACGTCCCAGGTCATTCTGAGGCCTGGCCCCAGTCCCAA 5249  
|||||  
5197 ggaatcagggagacgtcccaggtcattctgaggcctggccccagtcccaa 5246  
5250 CCCAACTGCTGTGCAGAATGAAAACCTGAAGAGCATGGCGCACAAGCGAA 5299  
|||||  
5247 cccaactgctgtgcagaatgaaaacctgaagagcatggcgcacaaagcgaa 5296  
5300 GCCAGCGCTCCAGTTACACAAGGCTCTCGAAAGATGCATCTGAGCTACAT 5349  
|||||  
5297 gccagcgctccagttacacaaggctctcgaaagatgcatctgagctacat 5346

[illegible]

Figure 24 (page 11 of 12)

5947 gatgagaaagtgagtttagcagaacccaaagccgtgctgggcccgcggtgac 5996  
|||||  
5947 gatgagaaagtgagtttagcagaacccaaagccgtgctgggcccgcggtgac 5996  
5997 tcattagcagaggaggagggggaggggcagtatattcctgggatactctctc 6046  
|||||  
5997 tcattagcagaggaggagggggaggggcagtatattcctgggatactctctc 6046  
6047 cagacccagcctggcttctgacatcatccacctgtgccctcaaaaccgtc 6096  
|||||  
6047 cagacccagcctggcttctgacatcatccacctgtgccctcaaaaccgtc 6096  
6097 ttagtctgttctgcaactcttaagtacatacctaactcagctcatggct 6146  
|||||  
6097 ttagtctgttctgcaactcttaagtacatacctaactcagctcatggct 6146  
6147 aaggaaaaaaaattaaagttgtcctgggtgattaaactctggacctccac 6196  
|||||  
6147 aaggaaaaaaaattaaagttgtcctgggtgattaaactctggacctccac 6196  
6197 atctaagtcctcgagttgacaaactgcacccccagcccaaacatctgttct 6246  
|||||  
6197 atctaagtcctcgagttgacaaactgcacccccagcccaaacatctgttct 6246  
6247 tccagagatctggaaagacatgggcataagagtcagtaaaactcttaatg 6296  
|||||  
6247 tccagagatctggaaagacatgggcataagagtcagtaaaactcttaatg 6296  
6297 aagggtgtcgggaagtgtttcccttaaagggggttagtgacagtacactag 6346  
|||||  
6297 aagggtgtcgggaagtgtttcccttaaagggggttagtgacagtacactag 6346  
6347 acaagtacagaatttagtttgctgacattaaatcactgccttctgggcc 6396  
|||||  
6347 acaagtacagaatttagtttgctgacattaaatcactgccttctgggcc 6396  
6397 gcaatagtttagtatctgttcctaaatcagcatttgagttctgcgggggt 6446  
|||||  
6397 gcaatagtttagtatctgttcctaaatcagcatttgagttctgcgggggt 6446  
6447 tacaaaatggaggcatgggttttaagaaggaatttgaaagcatccaggga 6496  
|||||  
6447 tacaaaatggaggcatgggttttaagaaggaatttgaaagcatccaggga 6496  
6497 gcaagaacagagtcggagctggtgtgaccatgacacttccattattgtt 6546  
|||||  
6497 gcaagaacagagtcggagctggtgtgaccatgacacttccattattgtt 6546

## Figure 24 (page 12 of 12)

6547 agcaaagctaggagaatgaccacatacattttaagtatcatgtaggctaga 6596  
|||||  
6547 agcaaagctaggagaatgaccacatacattttaagtatcatgtaggctaga 6596  
6597 gcgtgcctgtcatcatcatgcattgagaccgtttgtcctttgattatagt 6646  
|||||  
6597 gcgtgcctgtcatcatcatgcattgagaccgtttgtcctttgattatagt 6646  
6647 gtgtttatgtagaccagtcctagcagtatgtcctgagccatgggcaggat 6696  
|||||  
6647 gtgtttatgtagaccagtcctagcagtatgtcctgagccatgggcaggat 6696  
6697 tgtagatggcgaggtggcaggtttagtgtcgttgagatgtccatctctg 6746  
|||||  
6697 tgtagatggcgaggtggcaggtttagtgtcgttgagatgtccatctctg 6746  
6747 actggacctgtgtggtttactttagctgtcctcgtaaacactgccttgc 6796  
|||||  
6747 actggacctgtgtggtttactttagctgtcctcgtaaacactgccttgc 6796  
6797 tgtttgagtgttaattgagcattttcatgaagtgtatatgttgcatgatct 6846  
|||||  
6797 tgtttgagtgttaattgagcattttcatgaagtgtatatgttgcatgatct 6846  
6847 ttaacagggtatttttagacattaagaacagcatgatattactgacatatg 6896  
|||||  
6847 ttaacagggtatttttagacattaagaacagcatgatattactgacatatg 6896  
6897 cattagtagtatgcagtatcatgtaatgggtacttcagaggcagtttgg 6946  
|||||  
6897 cattagtagtatgcagtatcatgtaatgggtacttcagaggcagtttgg 6946  
6947 ttctttccagggtgagtagttatagctcacagtcattgggaatcctttattc 6996  
|||||  
6947 ttctttccagggtgagtagttatagctcacagtcattgggaatcctttattc 6996  
6997 acactgtctgtgaagtcgatgctggagtgccaaatgcacacggagactag 7046  
|||||  
6997 acactgtctgtgaagtcgatgctggagtgccaaatgcacacggagactag 7046  
7047 gaacaataaagacttgcatgccattctgactcatggagttttattcaata 7096  
|||||  
7047 gaacaataaagacttgcatgccattctgactcatggagttttattcaata 7096  
7097 gagcaatgataattccattctgattaaaaattaaaacacggaa 7140  
|||||  
7097 gagcaatgataattccattctgattaaaaattaaaacacggaa 7140

1	MSVLISQSVINYVEEENIPALKALLEKCKDVDERNECGQTPLMLAAEQGN	50
1	MSVLISQSVINYVEEENIPALKALLEKCKDVDERNECGQTPLMLAAEQGN	50
51	VEIVKELLKNGANCNLEDLDNWTALISASKEGHIHIVEELLKSGASLEHR	100
51	VEIVKELLKNGANCNLEDLDNWTALISASKEGHIHIVEELLKSGASLEHR	100
101	DMGGWTALMWACYKGRTDVVELLLSHGANPSVTGLQYSVYPIIWAAGRGH	150
101	DMGGWTALMWACYKGRTDVVELLLSHGANPSVTGL.YSVYPIIWAAGRGH	149
151	ADIVHLLLQNGAKVNCSDKYGTTPLVWAARKGHLECVKHLLAMGADV DQE	200
150	ADIVHLLLQNGAKVNCSDKYGTTPLVWAARKGHLECVKHLLAMGADV DQE	199
201	GANSTTALIVAVKGGYTQSVKEILKRNPVNLTDKDGNTALMIASKEGHI	250
200	GANSMTALIVAVKGGYTQSVKEILKRNPVNLTDKDGNTALMIASKEGHI	249
251	EIVQDLLDAGTYVNIPDRSGDTVLI GAVRGGHVEIVRALLQKYADIDIRG	300
250	EIVQDLLDAGTYVNIPDRSGDTVLI GAVRGGHVEIVRALLQKYADIDIRG	299
301	QDNKTALYWAVEKGNATMVRDILQCNPDTEICTKDGETPLIKATKMRNTE	350
300	QDNKTALYWAVEKGNATMVRDILQCNPDTEICTKDGETPLIKATKMRNIE	349
351	VVELLLDKGAKVSAVDKKGDTPLHV AIRGRSRRLAELLRLNPKDGRLLYR	400
350	VVELLLDKGAKVSAVDKKGDTPLHV AIRGRSRRLAELLRLNPKDGRLLYR	399
401	PNKAGETPYNIDCSHQKSILTQIFGARHLSPTETDGDMLGYDLYSSALAD	450
400	PNKAGETPYNIDCSHQKSILTQIFGARHLSPTETDGDMLGYDLYSSALAD	449
451	ILSEPTMQPPICVGLYAQWGS GKSFLLKKLEDEMKT FAGQQTEPLFQFSW	500
450	ILSEPTMQPPICVGLYAQWGS GKSFLLKKLEDEMKT FAGQQTEPLFQFSW	499
501	LIVFLTLLLCGGLGLVFAFTVDTNLAI AISLSFLALIYIFFIV IYFGGRR	550
500	LIVFLTLLLCGGLGLVFAFTVDTNLAI AISLSFLALIYIFFIV IYFGGRR	549

Figure 25 (page 2 of 4)

551 EGESWNWAWALSTRLARHIGYLELLFKLMFVNPPPELAEQTTKALPVRFLF 600  
|||||  
550 EGESWNWAWALSTRLARHIGYLELLFKLMFVNPPPELPEQTTKALPVRFLF 599  
601 TDYNRLSSVGGETSLAEMIATLSDACEREFGLATRLFRVLRTEESQGKK 650  
|||||  
600 TDYNRLSSVGGETSLAEMIATLSDACEREFGLATRLFRVFRTEESQGKK 649  
651 KWKRTCCLPSFVIFLFIVGCI IAGITLLAIFRVDPKHLTVNAILISIASV 700  
|||:|||||  
650 KWKRTCCLPSFVIFLFIVGCI IAGITLLAIFRVDPKHLTVNAILISIASV 699  
701 VGLAFVLNCRWWQVLDSSLNSQRKRLHSAASKLHKLKSEGFMKVLKCEV 750  
|||||  
700 VGLAFVLNCRWWQVLDSSLNSQRKRLHSAASKLHKLKSEGFMKVLKCEV 749  
751 ELMARMAKTIDSFTQNQTRLVVIIDGLDACEQDKVLQMLD TVRVLFSGKP 800  
|||||  
750 ELMARMAKTIDSFTQNQTRLVVIIDGLDACEQDKVLQMLD TVRVLFSGKP 799  
801 FIAIFASDPHIIKAINQNLNSVLRDSNINGHDYMRNIVHLPVFLNSRGL 850  
|||||  
800 FIAIFASDPHIIKAINQNLNSVLRDSNINGHDYMRNIVHLPVFLNSRGL 849  
851 SNARKFLVTSATNGDITCSDTTGTQEDTDRRVSONSLGEMTKLGSKTALN 900  
|||||  
850 SNARKFLVTSATNGDITCSDTTGTQEDTDRRVSONSLGEMTKLGSKTALN 899  
901 RRDYRRRQMORTITRQMSFDLTKLLVTEWFS DISPQTMRRLLNIVSVT 950  
|||||  
900 RRDYRRRQMORTITRQMSFDLTKLLVTEWFS DISPQTMRRLLNIVSVT 949  
951 GRLLRANQITFNWDRLASWINLTEQWPYRTSWLILYLEETEGLPDQMTLK 1000  
|||||  
950 GRLLRANQITFNWDRLASWINLTEQWPYRTSWLILYLEETEGLPDQMTLK 999  
1001 TIYERISKNIPTTKDVEPLLEIDGDIRNFEVFLSSRTPVLVARDVKTFP 1050  
|||||  
1000 TIYERISKNIPTTKDVEPLLEIDGDIRNFEVFLSSRTPVLVARDVKTFP 1049  
1051 CTVNLDPKLRETIADVRAAREQINIGGLAYPPLPLHEGPPRPPSGYSQPA 1100  
|||||  
1050 CTVNLDPKLREIIADVRAAREQINIGGLAYPPLPLHEGPPRPPSGYSQPA 1099  
1101 SVCSSASFNGPFPGGVVSPQPHSSYYSGLSGPQHFPFYNRPFFAPYLYTPR 1150  
|||||  
1100 SVCSSASFNGPFPGGVVSPQPHSSYYSGLSGPQHFPFYNRPFFAPYLYTPR 1149

1151	YYPGGSQHLISRSSVKASLPRDQNNGLPCDSGFNKQRQGSRPCRQLT	1200
1150	YYPGGSQHLISRSSVKTSLPRDQNNGLPCDSGFNKQRQA	1199
1201	EFNDRGCRCEKLRQIEGLDQSMMPQYCTTIKKANINGRVLSQCNIDELKK	1250
1200	SSMTVDVVC	1249
1251	EMAMNFGDWHLFRSMVLEMRSVESQVVPEDPRFLNENSSAPVPHGESARR	1300
1250	EMAMNFGDWHLFRSMVLEMRSVESQVVPEDPRFLNENSSAPVPHGESARR	1299
1301	PSHTELPLTELSSQTPYTLNFSFEELNTLGLDEGAPRHSNLSWQSQTRRT	1350
1300	SSHTELPLTELSSQTPYTLNFSFEELNTLGLDEGAPRHSNLSWQSQTRRT	1349
1351	PSLSSLNSQDSSIEISKLTDKVQAEYRDAYREYIAQMSQLEGGTGSSTIS	1400
1350	PSLSSLNSQDSSIEISKLTDKVQAEYRDAYREYIAQMSQLEGGTGSSTIS	1399
1401	GRSSPHSTYYIGQSSSGGSIHSTLEQERGKEGELKQEDGRKSFLMKRGDV	1450
1400	GRSSPHSTYYIGQSSSGGSIHSTLEQERGKEGELKQEDGRKSFLMKRGDV	1449
1451	IDYSSSGVSTNEASPLDPITEEDEKSDQSGSKLLPGKKSSERPSLFQTDL	1500
1450	IDYSSSGVSTNEASPLDPITEEDEKSDQSGSKLLPGKKSSERPSLFQTDL	1499
1501	KLKGGGLRYQKLPSDEDESGTEESDNTPLLKDDKDCKAEGKAERVCKSPE	1550
1500	KLKGGGLRYQKLPSDEDESGTEESDNTPLLKDDKDCKAEGKAERVCKSPE	1549
1551	HSAEPIRTFIKAKEYLSDALLDKKDSSDSGVRSNESSPNHSLHNEAADD	1600
1550	HSAEPIRTFIKAKEYLSDALLDKKDSSDSGVRSNESSPNHSLHNEAADD	1599
1601	QLEKANLIELEDEGHSGKRGMPHSLSGLDPIIARMSICSEDKKSPSECS	1650
1600	QLEKANLIELEDEGHSGKRGMPHSLSGLDPIIARMSICSEDKKSPSECS	1649
1651	LIASSPEESWPACQKAYNLNRTPTSTVTNLNNNTAPTNRANQNFDEIEGIRE	1700
1650	LIASSPEESWPACQKAYNLNRTPTSTVTNLNNNTAPTNRANQNFDEIEGIRE	1699
1701	TSQVILRPGPSNPPTAVQONENLKSMHKRSQRSSYTRL	1750
1700	TSQVILRPGPSNPPTAVQONENLKSMHKRSQRSSYTRL	1749

Figure 25 (page 4 of 4)

1751 ESTGFGEERESIL 1763

|||||||

1750 ESTGFGEERESIL 1762



(19) World Intellectual Property Organization  
International Bureau(43) International Publication Date  
14 March 2002 (14.03.2002)

PCT

(10) International Publication Number  
WO 02/020786 A3(51) International Patent Classification<sup>7</sup>: C12N 15/12,  
C07K 14/705, C12N 9/12, 15/63, 5/10, C07K 16/28,  
A61K 38/17, A01K 67/027, A61P 35/00Cancer Research Fund, 44 Lincoln's Inn Fields, London  
WC2A 3PX (GB). IGLESIAS, Teresa [ES/ES]; Insti-  
tuto de Investigaciones Biomédicas, CSIC, Calle Arturo  
Duperier, 4, E-28029 Madrid (ES).

(21) International Application Number: PCT/GB01/03977

(22) International Filing Date:  
6 September 2001 (06.09.2001)(74) Agent: MILES, John, S.; Eric Potter Clarkson, Park View  
House, 58 The Ropewalk, Nottingham NG1 5DD (GB).

(25) Filing Language: English

(26) Publication Language: English

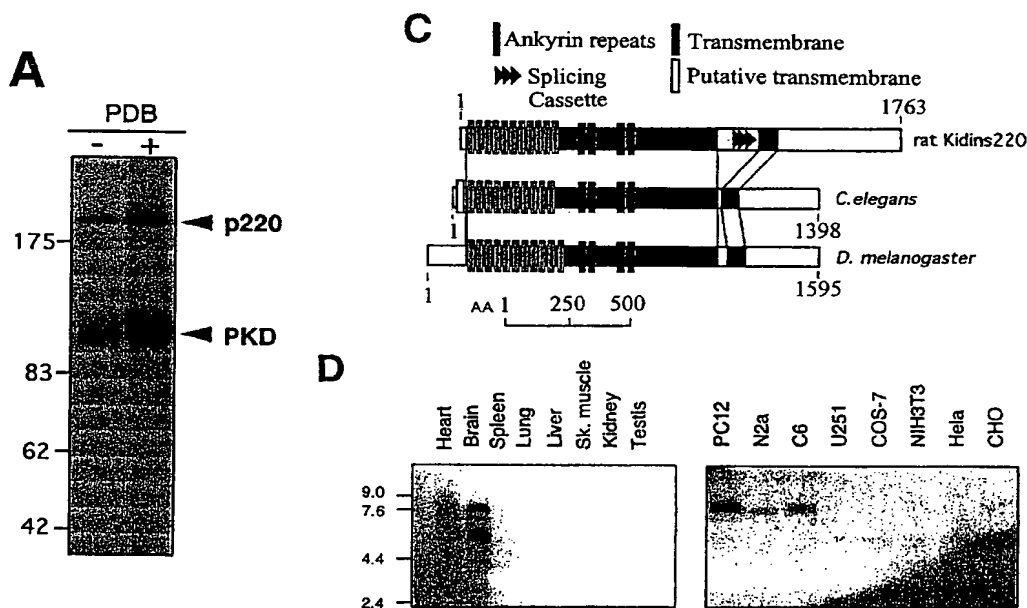
(30) Priority Data:  
60/230,449 6 September 2000 (06.09.2000) US(71) Applicant (for all designated States except US): IMPE-  
RIAL CANCER RESEARCH TECHNOLOGY LIM-  
ITED [GB/GB]; Sardinia House, Sardinia Street, London  
WC2A 3NL (GB).(81) Designated States (national): AE, AG, AL, AM, AT, AU,  
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU,  
CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH,  
GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC,  
LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW,  
MX, MZ, NO, NZ, PH, PL, PT, RO, RU, SD, SE, SG, SI,  
SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU,  
ZA, ZW.

(72) Inventors; and

(75) Inventors/Applicants (for US only): SCHIAVO, Giampi-  
etro [IT/GB]; Molecular Neuropathobiology, Imperial(84) Designated States (regional): ARIPO patent (GH, GM,  
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian  
patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European  
patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE,  
IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF,

[Continued on next page]

(54) Title: KINASE D INTERACTING PROTEIN



(57) Abstract: A polynucleotide comprising a nucleotide sequence which encodes the polypeptide whose amino acid sequence is given in Figure (1b) or Figure (11) or a fragment or variant or fusion thereof or a fusion of said fragment or variant, or its complement, the polypeptide encoded by said polynucleotide and fragments or variants or fusions thereof, antibodies specifically binding the said polypeptide or fragments or variants or fusions thereof. The invention further comprises methods of producing said polypeptide and antibodies and uses thereof.



CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

(88) Date of publication of the international search report:  
29 August 2002

**Published:**

— with international search report

*For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

## INTERNATIONAL SEARCH REPORT

International Application No  
PCT/GB 01/03977

## A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 C12N15/12 C07K14/705 C12N9/12 C12N15/63 C12N5/10  
C07K16/28 A61K38/17 A01K67/027 A61P35/00

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the International search (name of data base and, where practical, search terms used)

SEQUENCE SEARCH, CHEM ABS Data, EPO-Internal, WPI Data, PAJ, BIOSIS

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	DATABASE EM_HUM 'Online! EMBL; 11 November 1999 (1999-11-11) OHARA ET AL.: "Homo sapiens mRNA for KIAA1250 protein, partial cds." retrieved from EBI, accession no. AB033076 Database accession no. AB033076 XP002196566 cited in the application the whole document	1
A	-& DATABASE SWALL 'Online! 1 May 2000 (2000-05-01) NAGASE ET AL.: "KIAA1250 protein (fragment)" retrieved from EBI, accession no. Q9ULH0 Database accession no. Q9ULH0 XP002196567 cited in the application the whole document	1
A	-& NAGASE ET AL.: "Prediction of the -/-	1

☒ Further documents are listed in the continuation of box C.☐ Patent family members are listed in annex.

## \* Special categories of cited documents:

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

- "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- "&" document member of the same patent family

Date of the actual completion of the international search

18 April 2002

Date of mailing of the international search report

07/05/2002

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3016

Authorized officer

Ceder, O

## INTERNATIONAL SEARCH REPORT

International Application No

PCT/GB 01/03977

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
	coding sequences of unidentified human genes ..." DNA RESEARCH, vol. 6, 1999, pages 337-345, XP002943288 cited in the application abstract	
A	NISHIKAWA K ET AL: "DETERMINATION OF THE SPECIFIC SUBSTRATE SEQUENCE MOTIFS OF PROTEIN KINASE C ISOZYMES" JOURNAL OF BIOLOGICAL CHEMISTRY, AMERICAN SOCIETY OF BIOLOGICAL CHEMISTS, BALTIMORE, MD, US, vol. 272, no. 2, 10 January 1997 (1997-01-10), pages 952-960, XP002925439 ISSN: 0021-9258	
A	GSCHWENDT ET AL.: "Regulation of protein kinase Cmu by basic peptides and heparin" THE JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 272, no. 33, 15 August 1997 (1997-08-15), pages 20742-20746, XP002196563	
A	MATTHEWS ET AL.: "Characterization of serine 916 as an in vivo autophosphorylation site for protein kinase D/protein kinase Cmu" THE JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 274, no. 37, 10 September 1999 (1999-09-10), pages 26543-26549, XP002196564	
P,X	IGLESIAS ET AL.: "Identification and cloning of Kidins220, a novel neuronal substrate of protein kinase D" THE JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 275, no. 51, 22 December 2000 (2000-12-22), pages 40048-40056, XP002196565 the whole document	1-65

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

## FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

## Continuation of Box I.1

Although claims 98, 99 and 114 and claim 100 as far as it refers to claims 98-99 are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

Although claims 112-113 are directed to a diagnostic method practised on the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

---

## Continuation of Box I.1

Rule 39.1(iv) PCT - Method for treatment of the human or animal body by therapy

Rule 39.1(iv) PCT - Diagnostic method practised on the human or animal body

---

## Continuation of Box I.2

Present claims 67, 74-76, 79-81, 101, 102 and 104 relate to compounds/methods defined by reference to a desirable characteristic or property, namely

Claim 67 : method for modulating the function of the polypeptide of claims 33-42 by reducing the ability of said polypeptide to bind ATP;  
Claims 74- 76 and 79-81 (as far as they refer to claims 68-72): an agent identifiable by the method of claims 68-72;  
Claims 101 and 102: as far as they relate to an agent capable of inhibiting the function of the Kidins220 gene or its products, not being an antibody according to claim 51-61 or an antisense polynucleotide to the polynucleotides of claims 1-10, 17, 20, or 22;  
Claim 104: use of an agent which is capable of detecting the expression of the Kidins220 gene.

The claims cover all compounds/methods having the indicated characteristics or properties, whereas the application provides support within the meaning of Article 6 PCT and/or disclosure within the meaning of Article 5 PCT for only a very limited number of such products/compounds/methods/apparatus. In the present case, the claims so lack support, and the application so lacks disclosure, that a meaningful search over the whole of the claimed scope is impossible. Independent of the above reasoning, the claims also lack clarity (Article 6 PCT). An attempt is made to define the compounds/methods by reference to a result to be achieved. Again, this lack of clarity in the present case is such as to render a meaningful search over the whole of the claimed scope impossible. Consequently, the search has been carried out for those parts of the claims which appear to be clear, supported and disclosed, namely those parts relating to antibodies and antisense molecules of the claimed

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

polypeptides and polynucleotides.

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.